

**FIG. 1A**

SEQ ID NO: 1

PCMVII

```

→ 1  TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG
      AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC

      51  GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG
          CTCTGCCAGT GTCGAACAGA CATTGCGCTA CGGCCCTCGT CTGTTCTGGGC

      101  TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG
          AGTCCCGCGC AGTCGCCCAC AACCGCCCAC AGCCCCGACC GAATTGATAC

                                          HindIII
                                          -----

      151  CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTTGCA
          GCCGTAGTCT CGTCTAACAT GACTCTCACG TGGTATACTT CGAAAAACGT

      201  AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTAATTCTGG AATAGCTCAG
          TTTCGGATCC GGAGGTTTTT TCGGAGGAGT GATGAAGACC TTATCGAGTC

      251  AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA
          TCCGGCTCCG CCGGAGCCGG AGACGTATTT ATTTTTTTTA ATCAGTCGGT

      301  TGGGGCGGAG AATGGGCGGA ACTGGGCGGG GAGGGAATTA TTGGCTATTG
          ACCCCGCCTC TTACCCGCCT TGACCCGCCC CTCCTTAAT AACCGATAAC

      351  GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT
          CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

      401  CATGTCCAAT ATGACCGCCA TGTGACATT GATTATTGAC TAGTTATTAA
          GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT

      451  TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA TGGAGTCCG
          ATCATTAGTT AATGCCCCAG TAATCAAGTA TCGGGTATAT ACCTCAAGGC

      501  CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC
          GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG

      551  CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA
          GGGCGGGTAA CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT

      601  GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT AAAGTCCCCA
          CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA TTTGACGGGT

      651  CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCCC CCTATTGACG
          GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGCGGG GGATAACTGC

      701  TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCAGTA CATGACCTTA
          AGTTACTGCC ATTTACCGGG CGGACCGTAA TACGGGTCAT GTACTGGAAT

      751  CGGGACTTTC CTACTTGGA GTACATCTAC GTATTAGTCA TCGCTATTAC
          GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

      801  CATGGTGATG CGGTTTTGGC AGTACACCAA TGGGCGTGGA TAGCGGTTTG
          GTACCACTAC GCCAAAACCG TCATGTGGTT ACCCGCACCT ATCGCCAAAC

      851  ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG
          TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT AACTGCAGTT ACCCTCAAAC

```

FIG. 1B

901 TTTTGGCACC AAAATCAACG GGA CTTTCCA AAATGTCGTA ATAACCCCGC  
AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG

951 CCCGTTGACG CAAATGGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA  
GGGCAACTGC GTTACCCGC CATCCGCACA TGCCACCCTC CAGATATATT

1001 GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG CCATCCACGC  
CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC GGTAGGTGCG

1051 TGT TTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGCCG  
ACAAAACCTGG AGGTATCTTC TGTGGCCCTG GCTAGGTCTG AGGCGCCGGC

1101 GGAACGGTGC ATTGGAACGC GGATTCCCCG TGCCAAGAGT GACGTAAGTA  
CCTTGCCACG TAACCTTGCG CTAAGGGGC ACGGTTCTCA CTGCATTCTAT

1151 CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA  
GGCGGATATC TGAGATATCC GTGTGGGGAA ACCGAGAATA CGTACGATAT

1201 CTGT TTTTGG CTG GGGGCT ATACACCCCG GCTCCTTATG CTATAGGTGA  
GACAAAAACG GAACCCCGGA TATGTGGGGG CGAGGAATAC GATATCCACT

1251 TGGTATAGCT TAGCCTATAG GTGTGGGTTA TTGACCATTA TTGACCACTC  
ACCATATCGA ATCGGATATC CACACCCAAT AACTGGTAAT AACTGGTGAG

1301 CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG  
GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC

1351 CCACAACTAT CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT  
GGTGTGATA GAGATAACCG ATATACGGT ATGAGACAGG AAGTCTCTGA

1401 GACACGGACT CTGTATTTT ACAGGATGGG GTCCATTTAT TATTTACAAA  
CTGTGCCTGA GACATAAAAA TGTCTTACCC CAGGTAAATA ATAAATGTTT

1451 TTCACATATA CAACAACGCC GTCCCCCGTG CCCGCAGTTT TTATTAAACA  
AAGTGTATAT GTTGTGCGG CAGGGGGCAC GGGCGTCAA AATAATTTGT

1501 TAGCGTGGGA TCTCCGACAT CTCGGGTACG TGTTCGGAC ATGGGCTCTT  
ATCGCACCT AGAGGCTGTA GAGCCCATGC ACAAGGCCTG TACCCGAGAA

1551 CTCCGGTAGC GGCGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA  
GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGACCAGG GTAGGCAGGT

1601 GCGGCTCATG GTCGCTCGGC AGTCCTTGC TCCTAACAGT GGAGGCCAGA  
CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT

1651 CTTAGGCACA GCACAATGCC CACCACCACC AGTGTGCCGC ACAAGGCCGT  
GAATCCGTGT CGTGTACGG GTGGTGGTGG TCACACGGCG TGTTCGGCA

1701 GGCGGTAGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT  
CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGGA

1751 GGACGCAGAT GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT  
CCTGCGTCTA CCTTCTGAAT TCCGTCGCCG TCTTCTTCTA CGTCCGTCTA

1801 GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT TGCGGTGCTG  
CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA ACGCCACGAC

FIG. 1C

1851	TTAACGGTGG	AGGGCAGTGT	AGTCTGAGCA	GTACTCGTTG	CTGCCGCGCG
	AATTGCCACC	TCCCGTCACA	TCGACTCGT	CATGAGCAAC	GACGGCGCGC
1901	CGCCACCAGA	CATAATAGCT	GACAGACTAA	CAGACTGTTC	CTTTCCATGG
	GCGGTGGTCT	GTATTATCGA	CTGTCTGATT	GTCTGACAAG	GAAAGGTACC
			SalI	EcoRI	XhoI
			-----	-----	-----
1951	GTCTTTTCTG	CAGTCACCGT	CGTCGACCTA	AGAATTCAGA	CTCGAGCAAG
	CAGAAAAGAC	GTCAGTGGCA	GCAGCTGGAT	TCTTAAGTCT	GAGCTCGTTC
	XbaI	AscI	EcoRV	BamHI	MluI
	-----	-----	-----	-----	-----
2001	TCTAGAAAGG	CGCGCCAAGA	TATCAAGGAT	CCACTACGCG	TTAGAGCTCG
	AGATCTTTCC	GCGCGGTTCT	ATAGTTCCTA	GGTGATGCGC	AATCTCGAGC
2051	CTGATCAGCC	TCGACTGTGC	CTTCTAGTTG	CCAGCCATCT	GTTGTTTGCC
	GACTAGTCGG	AGCTGACACG	GAAGATCAAC	GGTCGGTAGA	CAACAAACGG
2101	CCTCCCCCGT	GCCTTCCTTG	ACCCTGGAAG	GTGCCACTCC	CACTGTCCTT
	GGAGGGGGCA	CGGAAGGAAC	TGGGACCTTC	CACGGTGAGG	GTGACAGGAA
2151	TCCTAATAAA	ATGAGGAAAT	TGCATCGCAT	TGTCTGAGTA	GGTGTCATTC
	AGGATTATTT	TACTCCTTTA	ACGTAGCGTA	ACAGACTCAT	CCACAGTAAG
2201	TATTCTGGGG	GGTGGGGTGG	GGCAGGACAG	CAAGGGGGAG	GATTGGGAAG
	ATAAGACCCC	CCACCCACC	CCGTCCGTGC	GTTCCCCCTC	CTAACCCCTC
2251	ACAATAGCAG	GCATGCTGGG	GAGCTCTTCC	GCTTCCTCGC	TCAGTGACTC
	TGTTATCGTC	CGTACGACCC	CTCGAGAAGG	CGAAGGAGCG	AGTGACTGAG
2301	GCTGCGCTCG	GTCGTTCGGC	TGCGGCGAGC	GGTATCAGCT	CACTCAAAGG
	CGACGCGAGC	CAGCAAGCCG	ACGCCGCTCG	CCATAGTCGA	GTGAGTTTCC
2351	CGGTAATACG	GTTATCCACA	GAATCAGGGG	ATAACGCAGG	AAAGAACATG
	GCCATTATGC	CAATAGGTGT	CTTAGTCCCC	TATTGCGTCC	TTTCTTGTA
2401	TGAGCAAAAAG	GCCAGCAAAA	GGCCAGGAAC	CGTAAAAAGG	CCGCGTTGCT
	ACTCGTTTTT	CGGTCGTTTT	CCGGTCCTTG	GCATTTTTTC	GGCGCAACGA
2451	GGCGTTTTTC	CATAGGCTCC	GCCCCCTGA	CGAGCATCAC	AAAAATCGAC
	CCGCAAAAAG	GTATCCGAGG	CGGGGGGACT	GCTCGTAGTG	TTTTTAGCTG
2501	GCTCAAGTCA	GAGGTGGCGA	AACCCGACAG	GACTATAAAG	ATACCAGGCG
	CGAGTTCAGT	CTCCACCGCT	TTGGGCTGTC	CTGATATTTT	TATGGTCCGC
2551	TTTCCCCCTG	GAAGCTCCCT	CGTGCGCTCT	CCTGTTCCGA	CCCTGCCGCT
	AAAGGGGGAC	CTTCGAGGGA	GCACGCGAGA	GGACAAGGCT	GGGACGGCGA
2601	TACCGGATAC	CTGTCCGCCT	TTCTCCCTTC	GGGAAGCGTG	GCGCTTTCTC
	ATGGCCTATG	GACAGGCGGA	AAGAGGGAAG	CCCTTCGCAC	CGCGAAAGAG
2651	AATGCTCACG	CTGTAGGTAT	CTCAGTTCGG	TGTAGGTCGT	TCGCTCCAAG
	TTACGAGTGC	GACATCCATA	GAGTCAAGCC	ACATCCAGCA	AGCGAGGTTT
2701	CTGGGCTGTG	TGCACGAACC	CCCCGTTTCA	CCCGACCGCT	GCGCCTTATC
	GACCCGACAC	ACGTGCTTGG	GGGGCAAGTC	GGGCTGGCGA	CGCGGAATAG

FIG. 1D

2751 CGGTAAGTAT CGTCTTGAGT CCAACCCGGT AAGACACGAC TTATCGCCAC  
GCCATTGATA GCAGAACTCA GGTTGGGCCA TTCTGTGCTG AATAGCGGTG

2801 TGGCAGCAGC CACTGGTAAC AGGATTAGCA GAGCGAGGTA TGTAGGCGGT  
ACCGTCGTCG GTGACCATTG TCCTAATCGT CTCGCTCCAT ACATCCGCCA

2851 GCTACAGAGT TCTTGAAGTG GTGGCCTAAC TACGGCTACA CTAGAAGGAC  
CGATGTCTCA AGAACTTCAC CACCGGATTG ATGCCGATGT GATCTTCCCTG

2901 AGTATTTGGT ATCTGCGCTC TGCTGAAGCC AGTTACCTTC GGAAAAAGAG  
TCATAAACCA TAGACGCGAG ACGACTTCGG TCAATGGAAG CCTTTTCTC

2951 TTGGTAGCTC TTGATCCGGC AAACAAACCA CCGCTGGTAG CGGTGGTTTT  
AACCATCGAG AACTAGGCCG TTTGTTTGGT GGCGACCATC GCCACCAAAA

3001 TTTGTTTGCA AGCAGCAGAT TACGCGCAGA AAAAAAGGAT CTCAAGAAGA  
AAACAAACGT TCGTCGTCTA ATGCGCGTCT TTTTTCCTA GAGTTCTTCT

3051 TCCTTTGATC TTTTCTACGG GGTCTGACGC TCAGTGGAAC GAAACTCAC  
AGGAAACTAG AAAAGATGCC CCAGACTGCG AGTCACCTTG CTTTGTAGTG

3101 GTTAAGGGAT TTTGGTCATG AGATTATCAA AAAGGATCTT CACCTAGATC  
CAATTCCCTA AAACCAGTAC TCTAATAGTT TTTCTAGAA GTGGATCTAG

3151 CTTTTAAATT AAAATGAAG TTTTAAATCA ATCTAAAGTA TATATGAGTA  
GAAAATTTAA TTTTACTTC AAAATTTAGT TAGATTTTCT ATATACTCAT

3201 AACTTGGTCT GACAGTTACC AATGCTTAAT CAGTGAGGCA CCTATCTCAG  
TTGAACCAGA CTGTCAATGG TTACGAATTA GTCACCTCGT GGATAGAGTC

3251 CGATCTGTCT ATTTTCGTTCA TCCATAGTTG CCTGACTCCC CGTCGTGTAG  
GCTAGACAGA TAAAGCAAGT AGGTATCAAC GGAAGTGGG GCAGCACATC

3301 ATAACACGA TACGGGAGGG CTTACCATCT GGCCCCAGTG CTGCAATGAT  
TATTGATGCT ATGCCCTCCC GAATGGTAGA CCGGGGTCAC GACGTTACTA

3351 ACCGCGAGAC CCACGCTCAC CGGCTCCAGA TTTATCAGCA ATAAACCAGC  
TGGCGCTCTG GGTGCGAGTG GCCGAGGTCT AAATAGTCGT TATTGGTCTG

3401 CAGCCGGAAG GGCCGAGCGC AGAAGTGGTC CTGCAACTTT ATCCGCCTCC  
GTCGGCCTTC CCGGCTCGCG TCTTCACCAG GACGTTGAAA TAGGCGGAGG

3451 ATCCAGTCTA TTAATTGTTG CCGGGAAGCT AGAGTAAGTA GTTCGCCAGT  
TAGGTCAGAT AATTAAACAAC GGCCCTTCGA TCTCATTCTA CAAGCGGTCA

3501 TAATAGTTTG CGCAACGTG TTGCCATTGC TACAGGCATC GTGGTGTCAC  
ATTATCAAAC GCGTTGCAAC AACGGTAACG ATGTCCGTAG CACCACAGTG

3551 GCTCGTCGTT TGGTATGGCT TCATTACGCT CCGGTTCCCA ACGATCAAGG  
CGAGCAGCAA ACCATACCGA AGTAAGTCGA GGCCAAGGGT TGCTAGTTCC

3601 CGAGTTACAT GATCCCCCAT GTTGTGCAAA AAAGCGGTTA GTCCTTCCG  
GCTCAATGTA CTAGGGGGTA CAACACGTTT TTTGCGCAAT CGAGGAAGCC

3651 TCCTCCGATC GTTGTGAGAA GTAAGTTGGC CGCAGTGTTA TCACTCATGG  
AGGAGGCTAG CAACAGTCTT CATTCAACCG GCGTCACAAT AGTGAGTACC

FIG. 1E

3701 TTATGGCAGC ACTGCATAAT TCTCTTACTG TCATGCCATC CGTAAGATGC  
AATACCGTCG TGACGTATTA AGAGAATGAC AGTACGGTAG GCATTCTACG

3751 TTTTCTGTGA CTGGTGAGTA CTCAACCAAG TCATTCTGAG AATAGTGTAT  
AAAAGACACT GACCACTCAT GAGTTGGTTC AGTAAGACTC TTATCACATA

3801 GCGGCGACCG AGTTGCTCTT GCGGCGGTC AATACGGGAT AATACCGCGC  
CGCCGCTGGC TCAACGAGAA CGGGCCGAG TTATGCCCTA TTATGGCGCG

3851 CACATAGCAG AACTTTAAAA GTGCTCATCA TTGAAAAACG TTCTTCGGGG  
GTGTATCGTC TTGAAATTTT CACGAGTAGT AACCTTTTGC AAGAAGCCCC

3901 CGAAAACTCT CAAGGATCTT ACCGCTGTTG AGATCCAGTT CGATGTAACC  
GCTTTTGAGA GTTCCTAGAA TGGCGACAAC TCTAGGTCAA GCTACATTGG

3951 CACTCGTGCA CCCAACTGAT CTTCAGCATC TTTTACTTTC ACCAGCGTTT  
GTGAGCACGT GGGTTGACTA GAAGTCGTAG AAAATGAAAG TGTCGCAAA

4001 CTGGGTGAGC AAAACAGGA AGGCAAAATG CCGCAAAAAA GGAATAAGG  
GACCCACTCG TTTTGTCTT TCCGTTTAC GCGTTTTTTT CCCTTATTCC

4051 GCGACACGGA AATGTTGAAT ACTCATACTC TTCCTTTTTC AATATTATTG  
CGCTGTGCCT TTACAACTTA TGAGTATGAG AAGGAAAAAG TTATAATAAC

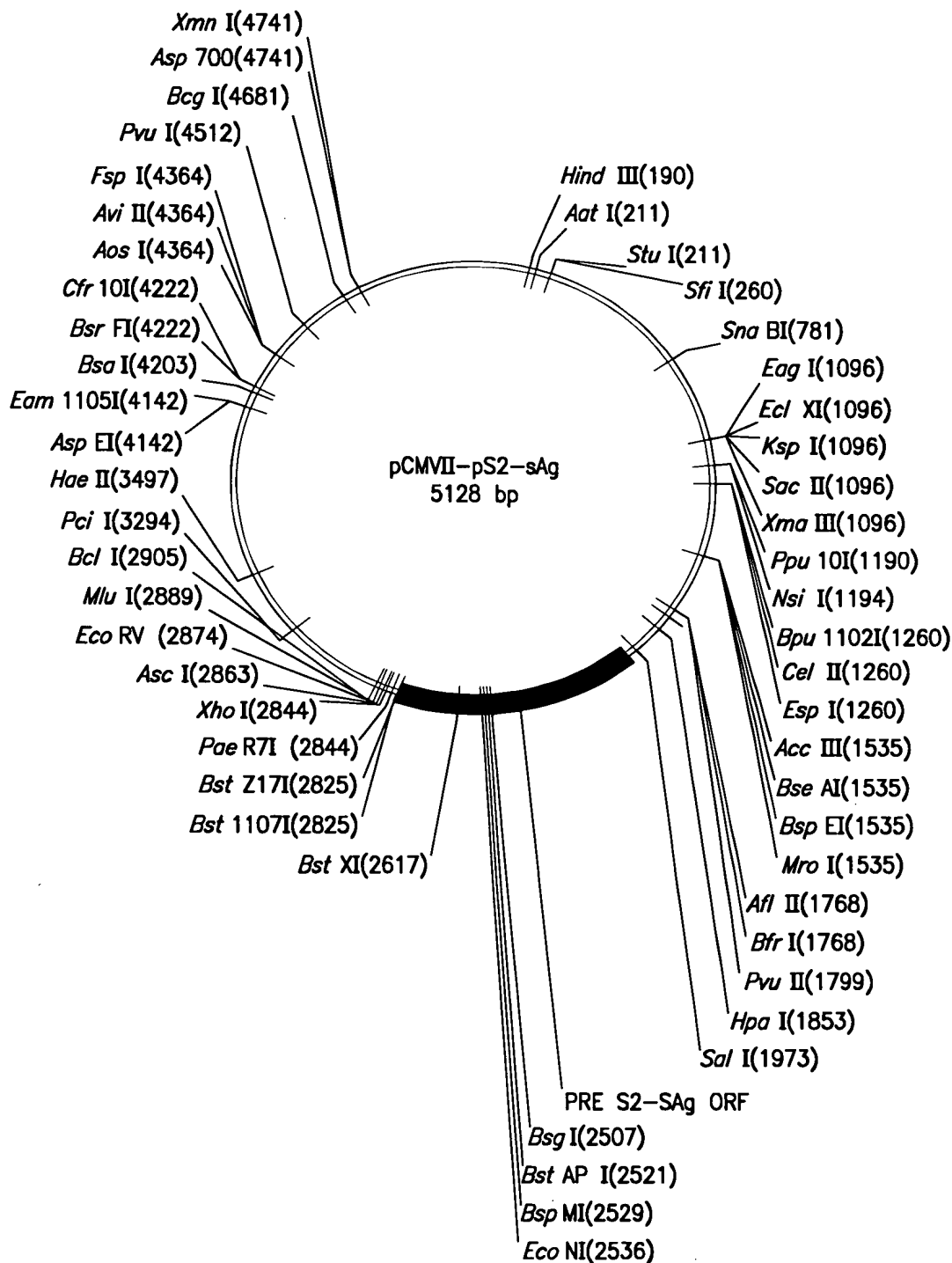
4101 AAGCATTIAT CAGGGTTATT GTCTCATGAG CGGATACATA TTTGAATGTA  
TTCGTAAATA GTCCCAATAA CAGAGTACTC GCCTATGTAT AAACCTACAT

4151 TTTAGAAAAA TAAACAAATA GGGGTTCCGC GCACATTTCC CCGAAAAGTG  
AAATCTTTTT ATTTGTTTAT CCCCAGGCG CGTGTAAGG GGCTTTTCAC

4201 CCACCTGACG TCTAAGAAAC CATTATTATC ATGACATTAA CCTATAAAAA  
GGTGGACTGC AGATTCTTTG GTAATAATAG TACTGTAATT GGATATTTTT

4251 TAGGCGTATC ACGAGGCCCT TTCGTC  
ATCCGCATAG TGCTCCGGA AAGCAG

FIG. 1F



**FIG. 2A**

SEQ ID NO: 2

```

→ 1  TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG
      AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC

      51  GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG
          CTCTGCCAGT GTCGAACAGA CATTCGCCCTA CGGCCCTCGT CTGTTTCGGGC

      101  TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG
          AGTCCCGCGC AGTCGCCAC  AACCGCCAC  AGCCCCGACC GAATTGATAC

                                          HindIII
                                          -----
      151  CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTTGCA
          GCCGTAGTCT CGTCTAACAT GACTCTCACG TGGTATACTT CGAAAAACGT

              StuI
              -----
              AatI
              -----
      201  AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTA CTCTCTGG AATAGCTCAG
          TTTCGGATCC GGAGGTTTTT TCGGAGGAGT GATGAAGACC TTATCGAGTC

              SfiI
              -----
      251  AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA
          TCCGGCTCCG CCGGAGCCGG AGACGTATTT ATTTTTTTTA ATCAGTCGGT

      301  TGGGGCGGAG AATGGGCGGA ACTGGGCGGG GAGGGAATTA TTGGCTATTG
          ACCCCGCCTC TTACCCGCCT TGACCCGCCC CTCCCTTAAT AACCGATAAC

      351  GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT
          CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

      401  CATGTCCAAT ATGACCGCCA TGTTGACATT GATTATTGAC TAGTTATTAA
          GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT

      451  TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA TGGAGTTCGG
          ATCATTAGTT AATGCCCCAG TAATCAAGTA TCGGGTATAT ACCTCAAGGC

      501  CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC
          GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG

      551  CCCGCCCAT TACGTCAATA ATGACGTATG TCCCATAGT AACGCCAATA
          GGGCGGGTAA CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT

      601  GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT AAAGTGCCCA
          CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA TTTGACGGGT

      651  CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCCC CCTATTGACG
          GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGCGGG GGATAACTGC

      701  TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCAGTA CATGACCTTA
          AGTTACTGCC ATTTACGGG CGGACCGTAA TACGGGTCAT GTACTGGAAT

                                          SnaBI
                                          -----
      751  CGGGACTTTC CTA CTCTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC
          GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

```

FIG. 2B



801	CATGGTGATG	CGGTTTTGGC	AGTACACCAA	TGGGCGTGGA	TAGCGGTTTG
	GTACCACTAC	GCCAAAACCG	TCATGTGGTT	ACCCGCACCT	ATCGCCAAAC
851	ACTCACGGGG	ATTTCCAAGT	CTCCACCCCA	TTGACGTCAA	TGGGAGTTTG
	TGAGTGCCCC	TAAAGGTTCA	GAGGTGGGGT	AACTGCAGTT	ACCCTCAAAC
901	TTTTGGCACC	AAAATCAACG	GGACTTTCCA	AAATGTCGTA	ATAACCCCGC
	AAAACCGTGG	TTTTAGTTGC	CCTGAAAGGT	TTTACAGCAT	TATTGGGGCG
951	CCCGTTGACG	CAAATGGGCG	GTAGGCGTGT	ACGGTGGGAG	GTCTATATAA
	GGGCAACTGC	GTTTACCCGC	CATCCGCACA	TGCCACCCTC	CAGATATATT
1001	GCAGAGCTCG	TTTAGTGAAC	CGTCAGATCG	CCTGGAGACG	CCATCCACGC
	CGTCTCGAGC	AAATCACTTG	GCAGTCTAGC	GGACCTCTGC	GGTAGGTGCG
				XmaIII	
				-----	
				SacII	
				-----	
				KspI	
				-----	
				EclXI	
				-----	
				EagI	
				-----	
1051	TGTTTTGACC	TCCATAGAAG	ACACCGGGAC	CGATCCAGCC	TCCGCGGCCG
	ACAAAACCTGG	AGGTATCTTC	TGTGGCCCTG	GCTAGGTCGG	AGGCGCCGGC
1101	GGAACGGTGC	ATTGGAACGC	GGATTCCCCG	TGCCAAGAGT	GACGTAAGTA
	CCTTGCCACG	TAACCTTGCG	CCTAAGGGGC	ACGGTTCTCA	CTGCATTTCAT
				Pp10I	
				-----	
				NsiI	
				-----	
1151	CCGCCTATAG	ACTCTATAGG	CACACCCCTT	TGGCTCTTAT	GCATGCTATA
	GGCGGATATC	TGAGATATCC	GTGTGGGGAA	ACCGAGAATA	CGTACGATAT
1201	CTGTTTTTGG	CTTGGGGCCT	ATACACCCCC	GCTCCTTATG	CTATAGGTGA
	GACAAAAACC	GAACCCCGGA	TATGTGGGGG	CGAGGAATAC	GATATCCACT
				EspI	
				-----	
				CelII	
				-----	
				Bpu1102I	
				-----	
1251	TGGTATAGCT	TAGCCTATAG	GTGTGGGTTA	TTGACCATTA	TTGACCACTC
	ACCATATCGA	ATCGGATATC	CACACCCAAT	AACTGGTAAT	AACTGGTGAG
1301	CCCTATTGGT	GACGATACTT	TCCATTACTA	ATCCATAACA	TGGCTCTTTG
	GGGATAACCA	CTGCTATGAA	AGGTAATGAT	TAGGTATTGT	ACCGAGAAAC
1351	CCACAACATAT	CTCTATTGGC	TATATGCCAA	TACTCTGTCC	TTCAGAGACT
	GGTGTTGATA	GAGATAACCG	ATATACGGTT	ATGAGACAGG	AAGTCTCTGA
1401	GACACGGACT	CTGTATTTTT	ACAGGATGGG	GTCCATTTAT	TATTTACAAA
	CTGTGCCTGA	GACATAAAAA	TGTCCTACCC	CAGGTAAATA	ATAAATGTTT

FIG. 2C

```

1451 TTCACATATA CAACAACGCC GTCCCCCGTG CCCGCAGTTT TTATTAAACA
    AAGTGTATAT GTTGTTGCGG CAGGGGGCAC GGGCGTCAAA AATAATTTGT

                                MroI
                                -----
                                BspEI
                                -----
                                BseAI
                                -----
                                AccIII
                                -----

1501 TAGCGTGGA TCTCCGACAT CTCGGGTACG TGTTCGGAC ATGGGCTCTT
    ATCGCACCCCT AGAGGCTGTA GAGCCCATGC ACAAGGCCTG TACCCGAGAA

1551 CTCCGGTAGC GCGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA
    GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT

1601 GCGGCTCATG GTCGCTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA
    CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT

1651 CTTAGGCACA GCACAATGCC CACCACCACC AGTGTGCCGC ACAAGGCCGT
    GAATCCGTGT CGTGTTACGG GTGGTGGTGG TCACACGGCG TGTTCCGGCA

1701 GGCGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT
    CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGGA

                                BfrI
                                -----
                                AflIII
                                -----
                                PvuII
                                -----

1751 GGACGCAGAT GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT
    CCTGCGTCTA CTTTCTGAAT TCCGTCGCCG TCTTCTTCTA CGTCCGTCTGA

                                PvuII
                                -----
                                HpaI
                                -----

1801 GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT TGCGGTGCTG
    CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA ACGCCACGAC

                                HpaI
                                -----

1851 TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG
    AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC

1901 CGCCACCAGA CATAATAGCT GACAGACTAA CAGACTGTTC CTTTCCATGG
    GCGGTGGTCT GTATTATCGA CTGTCTGATT GTCTGACAAG GAAAGGTACC

+2                               SEQ ID NO: 3—>M   Q   W   N
                                Sali
                                -----

1951 GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCATG CAGTGGAAC
    CAGAAAAGAC GTCAGTGGCA GCAGCTGGAT TCTTAAGTAT GTCACCTTGA

+2 S   T   A   F   H   Q   T   L   Q   D   P   R   V   R   G   L   Y
2001 CCACTGCCTT CCACCAAAC CTGCAGGATC CCAGAGTCAG GGGTCTGTAT
    GGTGACGGAA GGTGGTTTGA GACGTCCTAG GGTCTCAGTC CCCAGACATA

+2 L   P   A   G   G   S   S   S   G   T   V   N   P   A   P   N   I
2051 CTTCTGCTG GTGGCTCCAG TTCAGGAACA GTAAACCCTG CTCCGAATAT
    GAAGGACGAC CACCGAGGTC AAGTCCTTGT CATTTGGGAC GAGGCTTATA

```

FIG. 2D

```

+2  A S H I S S I S A R T G D P V T
2101 TGCTCTCAC ATCTCGTCAA TCTCCGCGAG GACTGGGGAC CCTGTGACGA
    ACGGAGAGTG TAGAGCAGTT AGAGGCGCTC CTGACCCCTG GGACACTGCT

+2  N M E N I T S G F L G P L L V L Q
2151 ACATGGAGAA CATCACATCA GGATTCTAG GACCCCTGCT CGTGTACAG
    TGTACCTCTT GTAGTGTAGT CCTAAGGATC CTGGGGACGA GCACAATGTC

+2  A G F F L L T R I L T I P Q S L D
2201 GCGGGGTTTT TCTTGTTGAC AAGAATCCTC ACAATACCGC AGAGTCTAGA
    CGCCCCAAAA AGAACAACTG TTCTTAGGAG TGTATGGCG TCTCAGATCT

+2  S W W T S L N F L G G S P V C L
2251 CTCGTGGTGG ACTTCTCTCA ATTTTCTAGG GGGATCTCCC GTGTGTCTTG
    GAGACCACC TGAAGAGAGT TAAAAGATCC CCCTAGAGGG CACACAGAAC

+2  G Q N S Q S P T S N H S P T S C P
2301 GCCAAAATTC GCAGTCCCA ACCTCCAATC ACTCACCAAC CTCCTGTCCT
    CGGTTTAAAG CGTCAGGGGT TGGAGGTTAG TGAGTGGTTG GAGGACAGGA

+2  P I C P G Y R W M C L R R F I I F
2351 CCAATTGTTC CTGGTTATCG CTGGATGTGT CTGCGGCGTT TTATCATATT
    GGTTAAACAG GACCAATAGC GACCTACACA GACGCCGCAA AATAGTATAA

+2  L F I L L L C L I F L L V L L D
2401 CCTCTTCATC CTGCTGCTAT GCCTCATCTT CTTATTGGTT CTTCTGGATT
    GGAGAAGTAG GACGACGATA CGGAGTAGAA GAATAACCAA GAAGACCTAA

+2  Y Q G M L P V C P L I P G S T T T
2451 ATCAAGGTAT GTTGCCCGTT TGTCTCTAA TTCCAGGATC AACACAACC
    TAGTTCCATA CAACGGGCAA ACAGGAGATT AAGGTCCTAG TTGTTGTTGG

+2  S T G P C K T C T T P A Q G N S M
    BstAPI
    -----
          BspMI          EcoNI
    -----
2501 AGTACGGGAC CATGCAAAAC CTGCACGACT CCTGCTCAAG GCAACTCTAT
    TCATGCCCTG GTACGTTTGT GACGTGCTGA GGACGAGTTC CGTTGAGATA
          BsgI
          -----

+2  F P S C C C T K P T D G N C T C
2551 GTTTCCCTCA TGTGCTGTA CAAAACCTAC GGATGGAAAT TGCACCTGTA
    CAAAGGGAGT ACAACGACAT GTTTTGGATG CCTACCTTTA ACGTGGACAT

+2  I P I P S S W A F A K Y L W E W A
    BstXI
    -----
2601 TTCCCATCCC ATCGTCCTGG GCTTTCGCAA AATACCTATG GGAGTGGGCC
    AAGGGTAGGG TAGCAGGACC CGAAAGCGTT TTATGGATAC CCTCACCCGG

+2  S V R F S W L S L L V P F V Q W F
2651 TCAGTCCGTT TCTCTGGCT CAGTTTACTA GTGCCATTTG TTCAGTGGTT
    AGTCAGGCAA AGAGAACCGA GTCAAATGAT CACGGTAAAC AAGTCACCAA

+2  V G L S P T V W L S A I W M M W
2701 CGTAGGGCTT TCCCCACTG TTTGGCTTTC AGCTATATGG ATGATGTGGT
    GCATCCCGAA AGGGGGTGAC AAACCGAAAG TCGATATACC TACTACACCA
  
```

FIG. 2E

Appln. No. 10/715,665  
Replacement Sheet

```

+2 Y W G P S L Y S I V S P F I P L L
2751 ATTGGGGGCC AAGTCTGTAC AGCATCGTGA GTCCCTTTAT ACCGCTGTTA
    TAACCCCGG TTCAGACATG TCGTAGCACT CAGGGAAATA TGGCGACAAT

+2 P I F F C L W V Y I *
                                BstZ17 I      XhoI
                                -----
                                Bst1107I     PaeR7I
                                -----

2801 CCAATTTTCT TTTGTCTCTG GGTATACATT TAAGAATTCA GACTCGAGCA
    GGTTAAAAGA AAACAGAGAC CCATATGTAA ATTCTTAAGT CTGAGCTCGT

                                AscI      EcoRV      MluI
                                -----

2851 AGTCTAGAAA GGCGCGCCAA GATATCAAGG ATCCACTACG CGTTAGAGCT
    TCAGATCTTT CCGCGCGGTT CTATAGTTCC TAGGTGATGC GCAATCTCGA

    BclI
    -----

2901 CGCTGATCAG CCTCGACTGT GCCTTCTAGT TGCCAGCCAT CTGTTGTTTG
    GCGACTAGTC GGAGCTGACA CGGAAGATCA ACGGTCGGTA GACAACAAAC

2951 CCCCTCCCCC GTGCCTTCCT TGACCCTGGA AGGTGCCACT CCCACTGTCC
    GGGGAGGGGG CACGGAAGGA ACTGGGACCT TCCACGGTGA GGGTGACAGG

3001 TTTCTTAATA AAATGAGGAA ATTGCATCGC ATTGTCTGAG TAGGTGTCAT
    AAAGGATTAT TTTACTCCTT TAACGTAGCG TAACAGACTC ATCCACAGTA

3051 TCTATTCTGG GGGGTGGGGT GGGGCAGGAC AGCAAGGGGG AGGATTGGGA
    AGATAAGACC CCCACCCCA CCGCTCCTG TCGTTCCCCC TCCTAACCCCT

3101 AGACAATAGC AGGCATGCTG GGGAGCTCTT CCGCTTCCTC GCTCACTGAC
    TCTGTTATCG TCCGTACGAC CCCTCGAGAA GGCGAAGGAG CGAGTGACTG

3151 TCGCTGCGCT CGGTCGTTCT GCTGCGGCGA GCGGTATCAG CTCACTCAAA
    AGCGACGCGA GCCAGCAAGC CGACGCCGCT CGCCATAGTC GAGTGAGTTT

                                Pci I
                                ---

3201 GGCGGTAATA CGGTTATCCA CAGAATCAGG GGATAACGCA GGAAAGAACA
    CCGCCATTAT GCCAATAGGT GTCTTAGTCC CCTATTGCGT CCTTTCCTGT

    Pci I
    ---

3251 TGTGAGCAAA AGGCCAGCAA AAGGCCAGGA ACCGTAAAAA GGCCGCGTTG
    AACTCGTTT TCCGGTCGTT TTCCGGTCCT TGGCATTTTT CCGGCGCAAC

3301 CTGGCGTTTT TCCATAGGCT CCGCCCCCT GACGAGCATC AAAAAATCG
    GACCGCAAAA AGGTATCCGA GCGGGGGGA CTGCTCGTAG TGTTTTTAGC

3351 ACGCTCAAGT CAGAGGTGGC GAAACCCGAC AGGACTATAA AGATACCAGG
    TGCGAGTTCA GTCTCCACCG CTTTGGGCTG TCCTGATATT TCTATGGTCC

3401 CGTTTCCCCC TGGAAGCTCC CTCGTGCGCT CTCCTGTTCC GACCCTGCCG
    GCAAAGGGGG ACCTTCGAGG GAGCACGCGA GAGGACAAGG CTGGGACGGC

```

FIG. 2F

					HaeII
					-----
3451	CTTACCGGAT	ACCTGTCCGC	CTTTCTCCCT	TCGGGAAGCG	TGGCGCTTTC
	GAATGGCCTA	TGGACAGGCG	GAAAGAGGGA	AGCCCTTCGC	ACCGCGAAAG
3501	TCAATGCTCA	CGCTGTAGGT	ATCTCAGTTC	GGTGTAGGTC	GTTTCGCTCCA
	AGTTACGAGT	GCGACATCCA	TAGAGTCAAG	CCACATCCAG	CAAGCGAGGT
3551	AGCTGGGCTG	TGTGCACGAA	CCCCCGTTC	AGCCCGACCG	CTGCGCCTTA
	TCGACCCGAC	ACACGTGCTT	GGGGGGCAAG	TCGGGCTGGC	GACGCGGAAT
3601	TCCGGTAACT	ATCGTCTTGA	GTCCAACCCG	GTAAGACACG	ACTTATCGCC
	AGGCCATTGA	TAGCAGAACT	CAGGTTGGGC	CATTCTGTGC	TGAATAGCGG
3651	ACTGGCAGCA	GCCACTGGTA	ACAGGATTAG	CAGAGCGAGG	TATGTAGGCG
	TGACCGTCGT	CGGTGACCAT	TGTCCTAATC	GTCTCGCTCC	ATACATCCGC
3701	GTGCTACAGA	GTTCTTGAAG	TGGTGGCCTA	ACTACGGCTA	CACTAGAAGG
	CACGATGTCT	CAAGAACTTC	ACCACCGGAT	TGATGCCGAT	GTGATCTTCC
3751	ACAGTATTTG	GTATCTGCGC	TCTGCTGAAG	CCAGTTACCT	TCGGAAAAAG
	TGTCATAAAC	CATAGACGCG	AGACGACTTC	GGTCAATGGA	AGCCTTTTTC
3801	AGTTGGTAGC	TCTTGATCCG	GCAAACAAAC	CACCGCTGGT	AGCGGTGGTT
	TCAACCATCG	AGAACTAGGC	CGTTTGTTTG	GTGGCGACCA	TCGCCACCAA
3851	TTTTTGTTTG	CAAGCAGCAG	ATTACGCGCA	GAAAAAAGG	ATCTCAAGAA
	AAAAACAAAC	GTTTCGTCGT	TAATGCGCGT	CTTTTTTTCC	TAGAGTTCTT
3901	GATCCTTTGA	TCTTTTCTAC	GGGGTCTGAC	GCTCAGTGGA	ACGAAAACTC
	CTAGGAAACT	AGAAAAGATG	CCCCAGACTG	CGAGTCACCT	TGCTTTTGAG
3951	ACGTTAAGGG	ATTTTGGTCA	TGAGATTATC	AAAAAGGATC	TTCACCTAGA
	TGCAATTCCC	TAAAACCAGT	ACTCTAATAG	TTTTTCCTAG	AAGTGGATCT
4001	TCCTTTTAAA	TTAAAAATGA	AGTTTTTAAAT	CAATCTAAAG	TATATATGAG
	AGGAAAATTT	AATTTTACT	TCAAAATTTA	GTTAGATTTC	ATATATACTC
4051	TAAACTTGGT	CTGACAGTTA	CCAATGCTTA	ATCAGTGAGG	CACCTATCTC
	ATTTGAACCA	GACTGTCAAT	GGTTACGAAT	TAGTCACTCC	GTGGATAGAG
					Eam1105I
					-----
					AspEI
					-----
4101	AGCGATCTGT	CTATTTTCGTT	CATCCATAGT	TGCCTGACTC	CCCCTCGTGT
	TCGCTAGACA	GATAAAGCAA	GTAGGTATCA	ACGGACTGAG	GGGCAGCACA
4151	AGATAACTAC	GATACGGGAG	GGCTTACCAT	CTGGCCCCAG	TGCTGCAATG
	TCTATTGATG	CTATGCCCTC	CCGAATGGTA	GACCGGGGTC	ACGACGTTAC

FIG. 2G

Cfr10I  
-----  
BsrFI  
-----

4201 ATACCGCGAG ACCCACGCTC ACCGGCTCCA GATTTATCAG CAATAAACCA  
TATGGCGCTC TGGGTGCGAG TGGCCGAGGT CTAAATAGTC GTTATTTGGT  
BsaI  
-----

4251 GCCAGCCGGA AGGGCCGAGC GCAGAAGTGG TCCTGCAACT TTATCCGCCT  
CGGTCGGCCT TCCCGGCTCG CGTCTTCACC AGGACGTTGA AATAGGCGGA

4301 CCATCCAGTC TATTAATTGT TGCCGGGAAG CTAGAGTAAG TAGTTCGCCA  
GGTAGGTCAG ATAATTAACA ACGGCCCTTC GATCTCATTC ATCAAGCGGT

FspI  
-----  
AviII  
-----  
AosI  
-----

4351 GTTAATAGTT TGCGCAACGT TGTTGCCATT GCTACAGGCA TCGTGGTGTG  
CAATTATCAA ACGCGTTGCA ACAACGGTAA CGATGTCCGT AGCACCACAG

4401 ACGCTCGTCG TTTGGTATGG CTTCAATTCAG CTCCGGTTCC CAACGATCAA  
TGCGAGCAGC AAACCATACC GAAGTAAGTC GAGGCCAAGG GTTGCTAGTT

4451 GGCAGATTAC ATGATCCCCC ATGTTGTGCA AAAAAGCGGT TAGCTCCTTC  
CCGCTCAATG TACTAGGGGG TACAACACGT TTTTTCGCCA ATCGAGGAAG

PvuI  
-----

4501 GGTCCTCCGA TCGTTGTCAG AAGTAAGTTG GCCGCAGTGT TATCACTCAT  
CCAGGAGGCT AGCAACAGTC TTCATTCAAC CGGCGTCACA ATAGTGAGTA

4551 GGTTATGGCA GCACTGCATA ATTCTCTTAC TGTCATGCCA TCCGTAAGAT  
CCAATACCGT CGTGACGTAT TAAGAGAATG ACAGTACGGT AGGCATTCTA

4601 GCTTTTCTGT GACTGGTGAG TACTCAACCA AGTCATTCTG AGAATAGTGT  
CGAAAAGACA CTGACCACTC ATGAGTTGGT TCAGTAAGAC TCTTATCACA

BcgI  
-----

4651 ATGCGGCGAC CGAGTTGCTC TTGCCCCGCG TCAATACGGG ATAATACCGC  
TACGCCGCTG GCTCAACGAG AACGGGCCGC AGTTATGCCC TATTATGGCG

XmnI  
-----  
Asp700  
-----

4701 GCCACATAGC AGAACTTTAA AAGTGCTCAT CATTGGAAAA CGTTCTTCGG  
CGGTGTATCG TCTTGAAATT TTCACGAGTA GTAACCTTTT GCAAGAAGCC

4751 GGCGAAAACCT CTCAAGGATC TTACCGCTGT TGAGATCCAG TTCGATGTAA  
CCGCTTTTGA GAGTTCCTAG AATGGCGACA ACTCTAGGTC AAGCTACATT

4801 CCCACTCGTG CACCCAACCTG ATCTTCAGCA TCTTTTACTT TCACCAGCGT  
GGGTGAGCAC GTGGGTTGAC TAGAAGTCGT AGAAAATGAA AGTGGTCGCA

FIG. 2H

4851 TTCTGGGTGA GCAAAAACAG GAAGGCAAAA TGCCGCAAAA AAGGGAATAA  
AAGACCCACT CGTTTTTGTC CTTCCTTTT ACGGCGTTTT TTCCCTTATT

4901 GGGCGACACG GAAATGTTGA ATACTCATAC TCTTCCTTTT TCAATATTAT  
CCCCTGTGTC CTTTACAAC TATGAGTATG AGAAGGAAAA AGTTATAATA

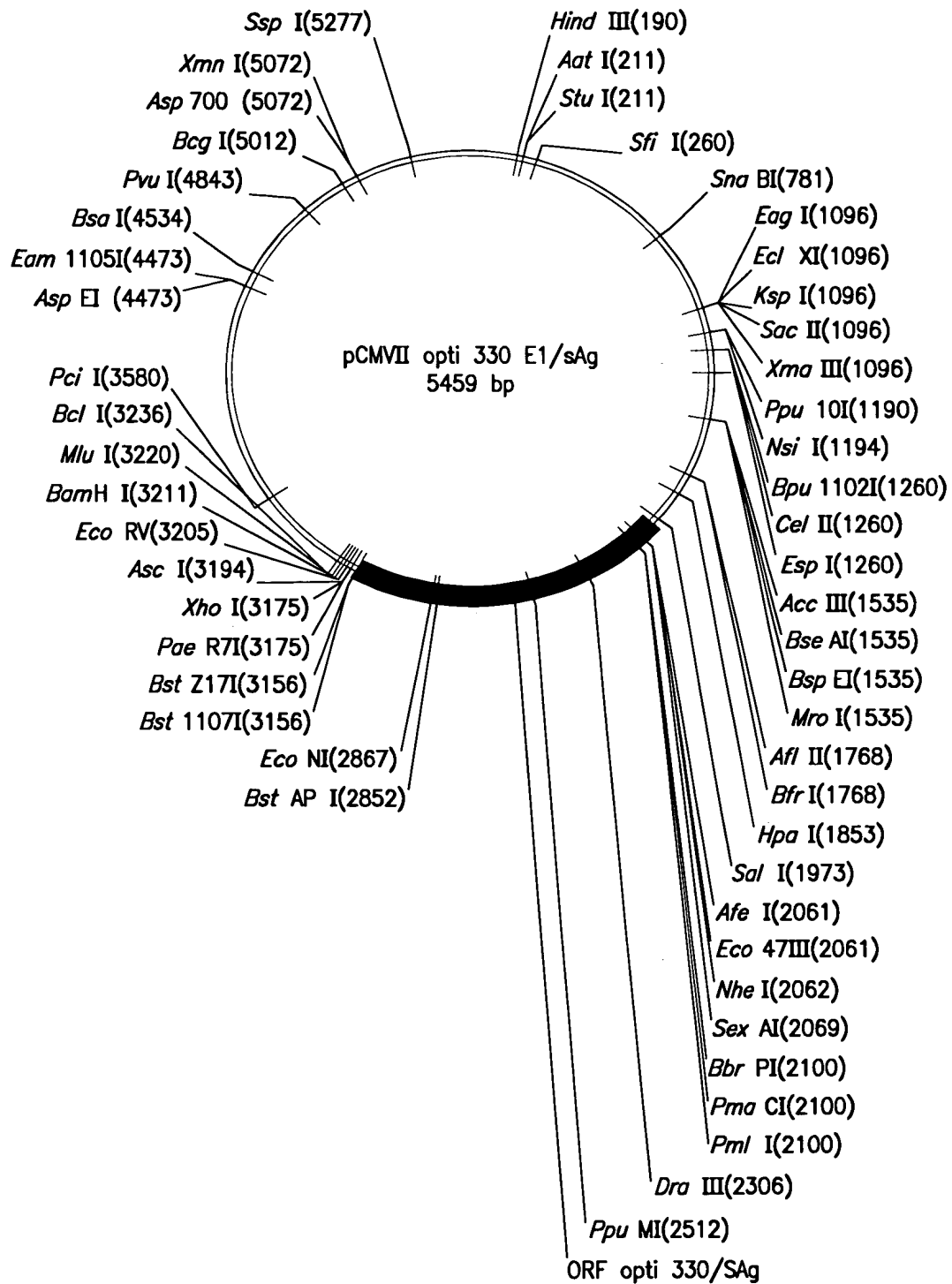
4951 TGAAGCATTT ATCAGGGTTA TTGTCTCATG AGCGGATACA TATTTGAATG  
ACTTCGTAAA TAGTCCAAT AACAGAGTAC TCGCCTATGT ATAAACTTAC

5001 TATTTAGAAA AATAAACAAA TAGGGGTTCC GCGCACATTT CCCCAGAAAAG  
ATAAATCTTT TTATTTGTTT ATCCCCAAGG CGCGTGTAAG GGGGCTTTTC

5051 TGCCACCTGA CGTCTAAGAA ACCATTATTA TCATGACATT AACCTATAAA  
ACGGTGGACT GCAGATTCTT TGGTAATAAT AGTACTGTAA TTGGATATTT

5101 AATAGGCGTA TCACGAGGCC CTTTCGTC  
TTATCCGCAT AGTGCTCCGG GAAAGCAG

FIG. 21



**FIG. 3A**



SEQ ID NO: 4

→ 1 TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG  
AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC

51 GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG  
CTCTGCCAGT GTCGAACAGA CATTTCGCTA CGGCCCTCGT CTGTTCTGGGC

101 TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG  
AGTCCCGCGC AGTCGCCCAC AACC GCCCAC AGCCCCGACC GAATTGATAC

HindIII  
-----  
151 CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTTGCA  
GCCGTAGTCT CGTCTAACAT GACTCTCACG TGGTATACTT CGAAAAACGT

StuI  
-----  
AatI  
-----  
201 AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG AATAGCTCAG  
TTTCGGATCC GGAGGTTTTT TCGGAGGAGT GATGAAGACC TTATCGAGTC

SfiI  
-----  
251 AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA  
TCCGGCTCCG CCGGAGCCGG AGACGTATTT ATTTTTTTTA ATCAGTCGGT

301 TGGGGCGGAG AATGGGCGGA ACTGGGCGGG GAGGGAATTA TTGGCTATTG  
ACCCCGCCTC TTACCCGCCT TGACCCGCCC CTCCCTTAAT AACC GATAAC

351 GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT  
CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

401 CATGTCCAAT ATGACCGCCA TGTGACATT GATTATTGAC TAGTTATTAA  
GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT

451 TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA TGGAGTCCG  
ATCATTAGTT AATGCCCCAG TAATCAAGTA TCGGGTATAT ACCTCAAGGC

501 CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC  
GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG

551 CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA  
GGGCGGGTAA CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT

601 GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT AAAGTCCCA  
CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA TTTGACGGGT

651 CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCCC CCTATTGACG  
GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGCGGG GGATAACTGC

701 TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCAGTA CATGACCTTA  
AGTTACTGCC ATTTACCGGG CGGACCGTAA TACGGGTCAT GTACTGGAAT

SnaBI  
-----  
751 CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC  
GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

FIG. 3B

801	CATGGTGATG	CGGTTTTGGC	AGTACACCAA	TGGGCGTGGA	TAGCGGTTTTG
	GTACCACTAC	GCCAAAACCG	TCATGTGGTT	ACCCGCACCT	ATCGCCAAAC
851	ACTCACGGGG	ATTTCCAAGT	CTCCACCCCA	TTGACGTCAA	TGGGAGTTTG
	TGAGTGCCCC	TAAAGGTTCA	GAGGTGGGGT	AACTGCAGTT	ACCCTCAAAC
901	TTTTGGCACC	AAAATCAACG	GGACTTTCCA	AAATGTCGTA	ATAACCCCGC
	AAAACCGTGG	TTTGTAGTTG	CCTGAAAGGT	TTTACAGCAT	TATTGGGGCG
951	CCCGTTGACG	CAAATGGGCG	GTAGGCGTGT	ACGGTGGGAG	GTCTATATAA
	GGGCAACTGC	GTTTACCCGC	CATCCGCACA	TGCCACCCTC	CAGATATATT
1001	GCAGAGCTCG	TTTAGTGAAC	CGTCAGATCG	CCTGGAGACG	CCATCCACGC
	CGTCTCGAGC	AAATCACTTG	GCAGTCTAGC	GGACCTCTGC	GGTAGGTGCG
				XmaIII	
				-----	
				SacII	
				-----	
				KspI	
				-----	
				EclXI	
				-----	
				EagI	
				-----	
1051	TGTTTTGACC	TCCATAGAAG	ACACCGGGAC	CGATCCAGCC	TCCGCGGCCG
	ACAAAACCTG	AGGTATCTTC	TGTGGCCCTG	GCTAGGTCGG	AGGCGCCCGG
1101	GGAACGGTGC	ATTGGAACGC	GGATTCCCCG	TGCCAAGAGT	GACGTAAGTA
	CCTTGCCACG	TAACCTTGCG	CCTAAGGGGC	ACGGTTCTCA	CTGCATTCTA
				PpuI 0I	
				-----	
				NsiI	
				-----	
1151	CCGCCTATAG	ACTCTATAGG	CACACCCCTT	TGGCTCTTAT	GCATGCTATA
	GGCGGATATC	TGAGATATCC	GTGTGGGGAA	ACCGAGAATA	CGTACGATAT
1201	CTGTTTTTTG	CTTGGGGCCT	ATACACCCCC	GCTCCTTATG	CTATAGGTGA
	GACAAAACCT	GAACCCCGGA	TATGTGGGGG	CGAGGAATAC	GATATCCACT
				EspI	
				-----	
				CelII	
				-----	
				Bpu1102I	
				-----	
1251	TGGTATAGCT	TAGCCTATAG	GTGTGGGTTA	TTGACCATTA	TTGACCACTC
	ACCATATCGA	ATCGGATATC	CACACCCAAT	AACTGGTAAT	AACTGGTGAG
1301	CCCTATTGGT	GACGATACTT	TCCATTACTA	ATCCATAACA	TGGCTCTTTG
	GGGATAACCA	CTGCTATGAA	AGGTAATGAT	TAGGTATTGT	ACCGAGAAAC
1351	CCACAACCTAT	CTCTATTGGC	TATATGCCAA	TACTCTGTCC	TTCAGAGACT
	GGTGTTGATA	GAGATAACCG	ATATACGGTT	ATGAGACAGG	AAGTCTCTGA
1401	GACACGGACT	CTGTATTTTT	ACAGGATGGG	GTCCATTTAT	TATTTACAAA
	CTGTGCCTGA	GACATAAAAA	TGTCCTACCC	CAGGTAAATA	ATAAATGTTT

FIG. 3C

1451 TTCACATATA CAACAACGCC GTCCCCCGTG CCCGCAGTTT TTATTAAACA  
AAGTGTATAT GTTGTTCGCG CAGGGGGCAC GGGCGTCAAA AATAATTTGT

MroI

-----

BspEI

-----

BseAI

-----

AccIII

-----

1501 TAGCGTGGGA TCTCCGACAT CTCGGGTACG TGTTCCGGAC ATGGGCTCTT  
ATCGCACCTT AGAGGCTGTA GAGCCCATGC ACAAGGCCTG TACCCGAGAA

1551 CTCCGGTAGC GGC GGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA  
GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CCGGACCAGG GTAGGCAGGT

1601 GCGGCTCATG GTCGCTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA  
CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT

1651 CTTAGGCACA GCACAATGCC CACCACCACC AGTGTGCCGC ACAAGGCCGT  
GAATCCGTGT CGTGTACCG GTGGTGGTGG TCACACGGCG TGTTCGGCA

1701 GGCGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT  
CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGGA

BfrI

-----

AflII

-----

1751 GGACGCAGAT GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT  
CCTGCGTCTA CCTTCTGAAT TCCGTCGCCG TCTTCTTCTA CGTCCGTCTGA

HpaI

-

1801 GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT TGCGGTGCTG  
CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA ACGCCACGAC

HpaI

-----

1851 TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG  
AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC

1901 CGCCACCAGA CATAATAGCT GACAGACTAA CAGACTGTTC CTTTCCATGG  
GCGGTGGTCT GTATTATCGA CTGTCTGATT GTCTGACAAG GAAAGGTACC

+3

SEQ ID NO: 5 —> M D A

SalI

-----

1951 GTCTTTTCTG CAGTCACCGT CGTCGACGAA TTCAAGCAAT CATGGATGCA  
CAGAAAAGAC GTCAGTGGCA GCAGCTGCTT AAGTTCGTTA GTACCTACGT

+3

2001 M K R G L C C V L L L C G A V F V  
ATGAAGAGAG GGCTCTGCTG TGTGCTGCTG CTGTGTGGAG CAGTCTTCGT  
TACTTCTCTC CCGAGACGAC ACACGACGAC GACACACCTC GTCAGAAGCA

FIG. 3D

```

+3   S P S   A S Y Q   V R N   S T G   L Y   H
      PmlI
      NheI
      -----
      Eco47III
      -----
      Afe I       SexAI
      -----
2051  TTCGCCCAGC GCTAGCTACC AGGTGCGCAA CAGCACCGGC CTGTACCACG
      AAGCGGGTCG CGATCGATGG TCCACGCGTT GTCGTGGCCG GACATGGTGC

+3   V T N D   C P N   S S I V   Y E A   A D A
      PmlI
      --
      PmaCI
      --
      BbrPI
      --
2101  TGACCAACGA CTGCCCCAAC AGCAGCATCG TGTACGAGGC CGCCGACGCC
      ACTGGTTGCT GACGGGGTTG TCGTCGTAGC ACATGCTCCG GCGGCTGCGG

+3   I L H T   P G C   V P C   V R E G   N A S
2151  ATCTGCACA CCCCCGGCTG CGTGCCCTGC GTGCGCGAGG GCAACGCCAG
      TAGGACGTGT GGGGGCCGAC GCACGGGACG CACGCGCTCC CGTTGCGGTC

+3   R C W   V A M T   P T V   A T R   D G K
2201  CCGCTGCTGG GTGGCCATGA CCCCCACCGT GGCCACCCGC GACGGCAAGC
      GCGGACGACC CACCGGTACT GGGGGTGGCA CCGGTGGGCG CTGCCGTTCG

+3   L P A T   Q L R   R H I D   L L V   G S A
      DraIII
2251  TGCCCCGCAC CCAGCTGCGC CGCCACATCG ACCTGCTGGT GGGCAGCGCC
      ACGGGCGGTG GGTCGACGCG GCGGTGTAGC TGGACGACCA CCCGTGCGGG

+3   T L C S   A L Y   V G D   L C G S   V F L
      DraIII
2301  ACCCTGTGCA GCGCCCTGTA CGTGGGCGAC CTGTGCGGCA GCGTGTTCCT
      TGGGACACGT CGCGGGACAT GCACCCGCTG GACACGCCGT CGCACAAGGA

+3   V G Q   L F T F   S P R   R H W   T T Q
2351  GGTGGGCCAG CTGTTACCT TCAGCCCCG CCGCCACTGG ACCACCCAGG
      CCACCCGGTC GACAAGTGA ATGCGGGGGC GGCGGTGACC TGGTGGGTCC

+3   G C N C   S I Y   P G H I   T G H   R M A
2401  GCTGCAACTG CAGCATCTAC CCCGGCCACA TCACCGGCCA CCGCATGGCC
      CGACGTTGAC GTCGTAGATG GGGCCGGTGT AGTGGCCGGT GGCGTACCGG

+3   W D M M   M N W   S P T   T M E N   I T S
2451  TGGGACATGA TGATGAACTG GAGCCCCACC ACCATGGAGA ACATCACATC
      ACCCTGTACT ACTACTTGAC CTCGGGGTGG TGGTACCTCT TGTAAGTGTAG

+3   G F L   G P L L   V L Q   A G F   F L L
      PpuMI
2501  AGGATTCCCTA GGACCCCTGC TCGTGTTACA GGCGGGGTTT TTCTTGTGTA
      TCCTAAGGAT CCTGGGGACG AGCACAATGT CCGCCCCAAA AAGAACAAC

```

FIG. 3E

Appln. No. 10/715,665  
Replacement Sheet

```

+3  T R I L   T I P   Q S L D   S W W   T S L
2551 CAAGAATCCT CACAATACCG CAGAGTCTAG ACTCGTGGTG GACTTCTCTC
    GTTCTTAGGA GTGTTATGGC GTCTCAGATC TGAGCACCAC CTGAAGAGAG

+3  N F L G   G S P   V C L   G Q N S   Q S P
2601 AATTTTCTAG GGGGATCTCC CGTGTGTCTT GGCCAAAATT CGCAGTCCCC
    TTAAAAGATC CCCCTAGAGG GCACACAGAA CCGGTTTTAA GCGTCAGGGG

+3  T S N   H S P T   S C P   P I C   P G Y
2651 AACCTCCAAT CACTCACCAA CCTCCTGTCC TCCAATTTGT CCTGGTTATC
    TTGGAGGTTA GTGAGTGGTT GGAGGACAGG AGGTAAACA GGACCAATAG

+3  R W M C   L R R   F I I F   L F I   L L L
2701 GCTGGATGTG TCTGCGGCGT TTTATCATAT TCCTCTTCAT CCTGCTGCTA
    CGACCTACAC AGACGCCGCA AAATAGTATA AGGAGAAGTA GGACGACGAT

+3  C L I F   L L V   L L D   Y Q G M   L P V
2751 TGCCTCATCT TCTTATTGGT TCTTCTGGAT TATCAAGGTA TGTTGCCCGT
    ACGGAGTAGA AGAATAACCA AGAAGACCTA ATAGTTCCAT ACAACGGGCA

+3  C P L   I P G S   T T T   S T G   P C K
                                     BstAPI
                                     -----
2801 TTGTCCTCTA ATTCCAGGAT CAACAACAAC CAGTACGGGA CCATGCAAAA
    AACAGGAGAT TAAGTCCTA GTTGTGTTG GTCATGCCCT GGTACGTTTT

+3  T C T T   P A Q   G N S M   F P S   C C C
    BstAP I           EcoNI
    -----
2851 CCTGCACGAC TCCTGCTCAA GGCAACTCTA TGTTTCCCTC ATGTTGCTGT
    GGACGTGCTG AGGACGAGTT CCGTTGAGAT ACAAAGGGAG TACAACGACA

+3  T K P T   D G N   C T C   I P I P   S S W
2901 ACAAACCTA CGGATGGAAA TTGCACCTGT ATCCCATCC CATCGTCTCTG
    TGTTTTGGAT GCCTACCTTT AACGTGGACA TAAGGGTAGG GTAGCAGGAC

+3  A F A   K Y L W   E W A   S V R   F S W
2951 GGCTTTTCGA AAATACCTAT GGGAGTGGGC CTCAGTCCGT TTCTCTTGGC
    CCGAAAGCGT TTTATGGATA CCCTCACCCG GAGTCAGGCA AAGAGAACCG

+3  L S L L   V P F   V Q W F   V G L   S P T
3001 TCAGTTTACT AGTGCCATTT GTTCAGTGGT TCGTAGGGCT TTCCCCACT
    AGTCAAATGA TCACGGTAAA CAAGTCACCA AGCATCCCGA AAGGGGGTGA

+3  V W L S   A I W   M M W   Y W G P   S L Y
3051 GTTTGGCTTT CAGCTATATG GATGATGTGG TATTGGGGGC CAAGTCTGTA
    CAAACCGAAA GTCGATATAC CTACTACACC ATAACCCCG GTTCAGACAT

+3  S I V   S P F I   P L L   P I F   F C L
3101 CAGCATCGTG AGTCCCTTTA TACCGCTGTT ACCAATTTTC TTTTGTCTCT
    GTCGTAGCAC TCAGGGAAAT ATGGCGACAA TGGTTAAAAG AAAACAGAGA

+3  W V Y I   *
    BstZ17 I           XhoI
    -----
    Bst1107I           PaeR7I           AscI
    -----
3151 GGGTATACAT TTAAGAATTC AGACTCGAGC AAGTCTAGAA AGGCGCGCCA
    CCCATATGTA AATTCTTAAG TCTGAGCTCG TTCAGATCTT TCCGCGCGGT

```

FIG. 3F

Appln. No. 10/715,665  
Replacement Sheet

	EcoRV	BamHI	MluI	BclI	
	-----	-----	-----	-----	
3201	AGATATCAAG TCTATAGTTC	GATCCACTAC CTAGGTGATG	GCGTTAGAGC CGCAATCTCG	TCGCTGATCA AGCGACTAGT	GCCTCGACTG CGGAGCTGAC
3251	TGCCTTCTAG ACGGAAGATC	TTGCCAGCCA AACGGTCGGT	TCTGTTGTTT AGACAACAAA	GCCCCTCCCC CGGGGAGGGG	CGTGCCTTCC GCACGGAAGG
3301	TTGACCCTGG AACTGGGACC	AAGGTGCCAC TTCCACGGTG	TCCCACTGTC AGGGTGACAG	CTTTCCTAAT GAAAGGATTA	AAAATGAGGA TTTTACTCCT
3351	AATTGCATCG TTAACGTAGC	CATTGTCTGA GTAACAGACT	GTAGGTGTCA CATCCACAGT	TTCTATTCTG AAGATAAGAC	GGGGGTGGGG CCCCCACCCC
3401	TGGGGCAGGA ACCCCGTCCT	CAGCAAGGGG GTCGTTCCCC	GAGGATTGGG CTCCTAACCC	AAGACAATAG TTCTGTTATC	CAGGCATGCT GTCCGTACGA
3451	GGGGAGCTCT CCCCTCGAGA	TCCGCTTCCT AGGCGAAGGA	CGCTCACTGA GCGAGTGACT	CTCGCTGCGC GAGCGACGCG	TCGGTCGTTC AGCCAGCAAG
3501	GGCTGCGGCG CCGACGCCGC	AGCGGTATCA TCGCCATAGT	GCTCACTCAA CGAGTGAGTT	AGGCGGTAAT TCCGCCATTA	ACGGTTATCC TGCCAATAGG
			Pci I -----		
3551	ACAGAATCAG TGTCTTAGTC	GGGATAACGC CCCTATTGCG	AGGAAAGAAC TCCTTTCTTG	ATGTGAGCAA TACACTCGTT	AAGGCCAGCA TTCCGGTCGT
3601	AAAGGCCAGG TTTCCGGTCC	AACCGTAAAA TTGGCATTTC	AGGCCGCGTT TCCGCGCAA	GCTGGCGTTT CGACCGCAA	TTCCATAGGC AAGGTATCCG
3651	TCCGCCCCC AGGCGGGGG	TGACGAGCAT ACTGCTCGTA	CACAAAAATC GTGTTTTTAG	GACGCTCAAG CTGCGAGTTC	TCAGAGGTGG AGTCTCCACC
3701	CGAAACCCGA GCTTTGGGCT	CAGGACTATA GTCCTGATAT	AAGATACCAG TTCTATGGTC	GCGTTTCCCC CGCAAAGGGG	CTGGAAGCTC GACCTTCGAG
3751	CCTCGTGCGC GGAGCACGCG	TCTCCTGTTC AGAGGACAAG	CGACCCTGCC GCTGGGACGG	GCTTACCGGA CGAATGGCCT	TACCTGTCCG ATGGACAGGC
3801	CCTTTCTCCC GGAAAGAGGG	TTCGGGAAGC AAGCCCTTCG	GTGGCGCTTT CACCGCGAAA	CTCAATGCTC GAGTTACGAG	ACGCTGTAGG TGCGACATCC
3851	TATCTCAGTT ATAGAGTCAA	CGGTGTAGGT GCCACATCCA	CGTTCGCTCC GCAAGCGAGG	AAGCTGGGCT TTCGACCCGA	GTGTGCACGA CACACGTGCT
3901	ACCCCCCGTT TGGGGGGCAA	CAGCCCGACC GTCGGGCTGG	GCTGCGCCTT CGACGCGGAA	ATCCGATAAC TAGGCCATTG	TATCGTCTTG ATAGCAGAAC
3951	AGTCCAACCC TCAGGTTGGG	GGTAAGACAC CCATTCTGTG	GACTTATCGC CTGAATAGCG	CACTGGCAGC GTGACCGTCG	AGCCACTGGT TCGGTGACCA
4001	AACAGGATTA TTGTCCTAAT	GCAGAGCGAG CGTCTCGCTC	GTATGTAGGC CATAATCCG	GGTGCTACAG CCACGATGTC	AGTTCTTGAA TCAAGAACTT
4051	GTGGTGGCCT CACCACCGGA	AACTACGGCT TTGATGCCGA	ACACTAGAAG TGTGATCTTC	GACAGTATTT CTGTCATAAA	GGTATCTGCG CCATAGACGC

FIG. 3G

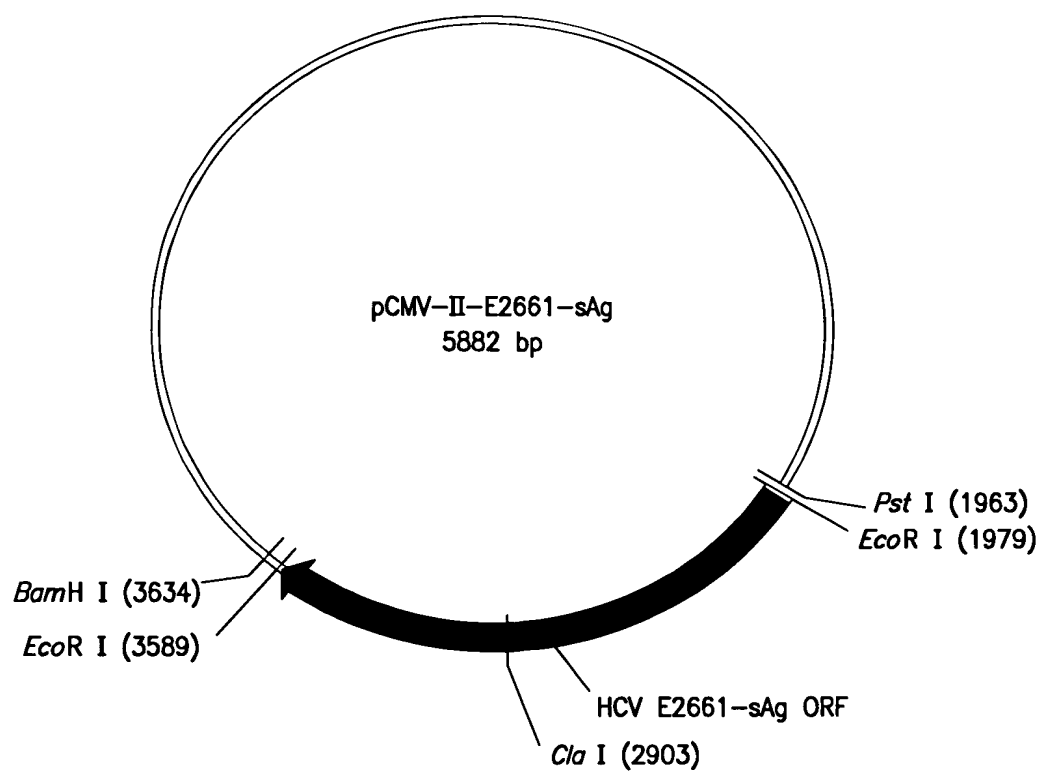
4101	CTCTGCTGAA	GCCAGTTACC	TTCGGAAAAA	GAGTTGGTAG	CTCTTGATCC
	GAGACGACTT	CGGTCAATGG	AAGCCTTTTT	CTCAACCATC	GAGAACTAGG
4151	GGCAAACAAA	CCACCGCTGG	TAGCGGTGGT	TTTTTTGTTT	GCAAGCAGCA
	CCGTTTGTTT	GGTGGCGACC	ATCGCCACCA	AAAAAACAAA	CGTTCGTCGT
4201	GATTACGCGC	AGAAAAAAG	GATCTCAAGA	AGATCCTTTG	ATCTTTTCTA
	CTAATGCGCG	TCTTTTTTTC	CTAGAGTTCT	TCTAGGAAAC	TAGAAAAGAT
4251	CGGGGTCTGA	CGCTCAGTGG	AACGAAAACT	CACGTTAAGG	GATTTTGGTC
	GCCCCAGACT	GCGAGTCACC	TTGCTTTTGA	GTGCAATTCC	CTAAAACCAG
4301	ATGAGATTAT	CAAAAAGGAT	CTTCACCTAG	ATCCTTTTAA	ATTAAAAATG
	TACTCTAATA	GTTTTTCCTA	GAAGTGGATC	TAGGAAAATT	TAATTTTATC
4351	AAGTTTTTAAA	TCAATCTAAA	GTATATATGA	GTAAACTTGG	TCTGACAGTT
	TTCAAAATTT	AGTTAGATT	CATATATACT	CATTTGAACC	AGACTGTCAA
4401	ACCAATGCCT	AATCAGTGAG	GCACCTATCT	CAGCGATCTG	TCTATTTTCGT
	TGGTTACGAA	TTAGTCACTC	CGTGGATAGA	GTCGCTAGAC	AGATAAAGCA
Eam1105I					
-----					
AspEI					
-----					
4451	TCATCCATAG	TTGCCTGACT	CCCCGTCGTG	TAGATAACTA	CGATACGGGA
	AGTAGGTATC	AACGGACTGA	GGGGCAGCAC	ATCTATTGAT	GCTATGCCCT
4501	GGGCTTACCA	TCTGGCCCCA	GTGCTGCAAT	GATACCGCGA	GACCCACGCT
	CCCGAATGGT	AGACCGGGGT	CACGACGTTA	CTATGGCGCT	CTGGGTGCGA
BsaI					
-----					
4551	CACCGGCTCC	AGATTTATCA	GCAATAAACC	AGCCAGCCGG	AAGGGCCGAG
	GTGGCCGAGG	TCTAAATAGT	CGTTATTTGG	TCGGTCGGCC	TTCCCGGCTC
4601	CGCAGAAGTG	GTCCTGCAAC	TTTATCCGCC	TCCATCCAGT	CTATTAATTG
	GCGTCTTCAC	CAGGACGTTG	AAATAGGCGG	AGGTAGGTCA	GATAATTAAC
4651	TTGCCGGGAA	GCTAGAGTAA	GTAGTTCGCC	AGTTAATAGT	TTGCGCAACG
	AACGGCCCTT	CGATCTCATT	CATCAAGCGG	TCAATTATCA	AACGCGTTGC
4701	TTGTTGCCAT	TGCTACAGGC	ATCGTGGTGT	CACGCTCGTC	GTTTGGTATG
	AACAACGGTA	ACGATGTCCG	TAGCACCACA	GTGCGAGCAG	CAAACCATAC
4751	GCTTCATTCA	GCTCCGGTTC	CCAACGATCA	AGGCGAGTTA	CATGATCCCC
	CGAAGTAAGT	CGAGGCCAAG	GGTTGCTAGT	TCCGCTCAAT	GTACTAGGGG
PvuI					
-----					
4801	CATGTTGTGC	AAAAAAGCGG	TTAGCTCCTT	CGGTCCTCCG	ATCGTTGTCA
	GTACAACACG	TTTTTTCGCC	AATCGAGGAA	GCCAGGAGGC	TAGCAACAGT
4851	GAAGTAAGTT	GGCCGCAGTG	TTATCACTCA	TGGTTATGGC	AGCACTGCAT
	CTTCATTCAA	CCGGCGTCAC	AATAGTGAGT	ACCAATACCG	TCGTGACGTA
4901	AATTCTCTTA	CTGTCATGCC	ATCCGTAAGA	TGCTTTTCTG	TGACTGGTGA
	TTAAGAGAAT	GACAGTACGG	TAGGCATTCT	ACGAAAAGAC	ACTGACCACT

FIG. 3H

					BcgI
					-----
4951	GTACTCAACC	AAGTCATTCT	GAGAATAGTG	TATGCGGCGA	CCGAGTTGCT
	CATGAGTTGG	TTCAGTAAGA	CTCTTATCAC	ATACGCCGCT	GGCTCAACGA
5001	CTTGCCCGGC	GTCAATACGG	GATAATACCG	CGCCACATAG	CAGAACTTTA
	GAACGGGCCG	CAGTTATGCC	CTATTATGGC	GCGGTGTATC	GTCTTGAAAT
					XmnI
					-----
					Asp700
					-----
5051	AAAGTGCTCA	TCATTGGAAG	ACGTTCTTCG	GGGCGAAAAC	TCTCAAGGAT
	TTTCACGAGT	AGTAACCTTT	TGCAAGAAGC	CCCCTTTTGT	AGAGTTCCTA
5101	CTTACCGCTG	TTGAGATCCA	GTTCGATGTA	ACCCACTCGT	GCACCCAACT
	GAATGGCGAC	AAGTCTAGGT	CAAGCTACAT	TGGGTGAGCA	CGTGGGTTGA
5151	GATCTTCAGC	ATCTTTTACT	TTCACCAGCG	TTTCTGGGTG	AGCAAAAACA
	CTAGAAGTCG	TAGAAAATGA	AAGTGGTCGC	AAAGACCCAC	TCGTTTTTGT
5201	GGAAGGCAAA	ATGCCGCAAA	AAAGGGAATA	AGGGCGACAC	GGAATGTTG
	CCTTCCGTTT	TACGGCGTTT	TTTCCCTTAT	TCCCGCTGTG	CCTTTACAAC
					SspI
					-----
5251	AATACTCATA	CTCTTCCTTT	TTCAATATTA	TTGAAGCATT	TATCAGGGTT
	TTATGAGTAT	GAGAAGGAAA	AAGTTATAAT	AACTTCGTAA	ATAGTCCCAA
5301	ATTGTCTCAT	GAGCGGATAC	ATATTTGAAT	GTATTTAGAA	AAATAAACAA
	TAACAGAGTA	CTCGCCTATG	TATAAACTTA	CATAAATCTT	TTTATTTGTT
5351	ATAGGGGTTC	CGCGCACATT	TCCCCGAAAA	GTGCCACCTG	ACGTCTAAGA
	TATCCCCAAG	GCGCGTGTA	AGGGGCTTTT	CACGGTGGAC	TGCAGATTCT
5401	AACCATTATT	ATCATGACAT	TAACCTATAA	AAATAGGCGT	ATCACGAGGC
	TTGGTAATAA	TAGTACTGTA	ATTGGATATT	TTTATCCGCA	TAGTGCTCCG
5451	CCTTTCGTC				
	GGAAAGCAG				

FIG. 31





**FIG. 4A**

pcmv-II - E2661-sAg

SEQ ID NO: 6  
→ 1 TC GCGCGGTTT CGTGTATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGTCT GTAAGCGGAT  
AGCGGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGCG CTCTGCCAGT GTCGAACAGA CATTGCGCTA

81 GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACATATG CGGCATCAGA  
CGGCCCTCGT CTGTTGCGGC AGTCCCGCGC AGTCGCCAC AACCGCCAC AGCCCCGACC GAATTGATAC GCCGTAGTCT

161 GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTTTGA AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG  
CGTCTAACAT GACTCTCAG TGGTATACTT CGAAAAACGT TTTCGGATCC GGAGTTT TCGGAGGAGT GATGAAGACC

241 AATAGCTCAG AGGCCGAGGC GGCTCGGCC TCTGCATAAA TAAAAAAAT TAGTCAGCCA TGGGGCGGAG AATGGCGGA  
TTATCGAGTC TCCGGCTCCG CCGGAGCCGG AGACGTATTT ATTTTTTTTA ATCAGTCGGT ACCCCGCTC TTACCCGCT

321 ACTGGGCGGG GAGGGAATTA TTGGCTATTG GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT  
TGACCCGCCC CTCCTTAAT AACCGATAAC CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

401 CATGTCCAAT ATGACCGCCA TGTGACATT GATTATTGAC TAGTTAATAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT  
GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAGTA

481 AGCCCATATA TGGAGTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCCGCCCAT  
TCGGGTATAT ACCTCAAGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGSC GGGTTGTGG GGGCGGGTAA

561 GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTACGGT  
CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA

641 AACTGCCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCCC CCTATTGACG TCAATGACGG TAAATGGCCC  
TTTGACGGGT GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGCGGG GGATAACTGC AGTTACTGCC ATTTACCGGG

721 GCCTGGCATT ATGCCCAGTA CATGACCTTA CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC  
CGGACCGTAA TACGGGTCAAT GTACTGGAAT GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

FIG. 4B

801 CATGGTGATG CGGTTTTTGGC AGTACACCAG TGGCGGTGGA TAGCGTTTGG ACTCAGCGGG ATTTCCAAGT CTCCACCCCA  
GTACCACTAC GCCAAAACCG TCATGTGTTT ACCCGCACCT ATCGCCAAAC TGAGTGCCCC TAAAGGTCA GAGGTGGGT

881 TTGACGTCAA TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTTCCA AAATGTCGTA ATAAACCCGC CCGTTGACG  
AACTGCAGTT ACCCTCAAAC AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG GGGCAACTGC

961 CAAATGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG  
GTTTACCCGC CATCCGCACA TGCCACCCCTC CAGATATAAT CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC

1041 CCATCCACGC TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGGCG GGAACGGTGC ATTGGAACGC  
GGTAGGTGCG ACAAAACTGG AGGTATCTTC TGTGGCCCTG GCTAGGTGCG AGGCGCCGCG CCTTGCCACG TAACCTTGCG

1121 GGATTCCCC TGCCAAGAGT GACGTAAGTA CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA  
CCTAAGGGC ACGGTTCTCA CTGCATTTCAT GCGGATATC TGAGATATCC GTGTGGGAA ACCGAGAATA CGTACGATAT

1201 CTGTTTTTGG CTGTGGGCT ATACACCCCC GCTCCTTAG CTATAGTGA TGGTATAGCT TAGCCTATAG GGTGGGTTA  
GACAAAACC GAACCCCGGA TATGTGGGG CGAGGAATAC GATATCCACT ACCATATCGA ATCGGATATC CACACCAAT

1281 TTGACCATTA TTGACCACTC CCCTATTGGT GACGATACCT TCCATTACTA ATCCATAACA TGGCTCTTTG CCACAACTAT  
AACTGGTAAT AACTGGTGAG GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC GGTGTTGATA

1361 CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT GACACGACT CTGTATTTT ACAGGATGGG GTCCATTTAT  
GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA CTGTGCTGTA GACATAAAA TGTCCTACCC CAGGTAATA

1441 TATTTACAAA TTCACATATA CAACAACGCC GTCCCCCGTG CCGCAGTTT TTATTAAACA TAGCGTGGGA TCTCCGACAT  
ATAAATGTTT AAGTGATAT GTTGTGCGG CAGGGGGCAC GGGCGTCAAA AATAATTGT ATCGCACCT AGAGGCTGTA

1521 CTCGGGTACG TGTTCGGAC ATGGGCTCTT CTCCGGTAGC GCGGAGCTT CCACATCGA GCCCTGGTCC CATCGTCCA  
GAGCCCATGC ACAAGGCCTG TACCCGAGAA GAGGCCATCG CCGCCTCGAA GGTGTAGGCT GGGGACCAGG GTAGGCAGGT

FIG. 4C

```

1601 GCGGTCATG GTCGCTCGG AGCTCCTTGG TCCTAACAGT GGAGGCCAGA CTTAGGCACA GCACAATGCC CACCACCACC
CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT GAATCCGTGT CGTGTTACGG GTGGTGGTGG

1681 AGTGTGCCG ACAAGGCCGT GCGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT GGACGCAGAT
TCACACGGCG TGTTCGGCA CGCCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGGG CCTGCGTCTA

1761 GGAAGACTTA AGGAGCGGC AGAAGAAGAT GCAGGCAGCT GAGTTGTTGT ATTCTGATTA GAGTCAGAGG TAACTCCCGT
CCTTCTGAAT TCCGTCGCCG TCTTCTTCTA CGTCCGTCGA CTCAACAACA TAAGACTATT CGCAGTCTCC ATTGAGGGCA

1841 TCGGTGCTG TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG GCCCACCAGA CATAATAGCT
ACGCCACGAC AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC GCGGTGGTCT GTATTATCGA

+3      SEQ ID NO:7 -----> M D A
      PstI                      EcoRI
      -----
1921 GACAGACTAA CAGACTGTTT CTTTCCATGG GTCTTTTCTG CAGTCACCGT CGTCGACGAA TTCAAGCAAT CATGGATGCA
CTGTCTGATT GTCTGACAAG GAAAGGTACC CAGAAAGAC GTCAGTGGCA GCAGCTGCTT AAGTTCTGTTA GTACCTACGT

+3 M K R G L C C V L L L L C G A V F V S P S A S E T H V T
2001 ATGAAGAGAG GGCTCTGCTG TGTGCTGCTG CTGTGTGGAG CAGTCTTCTG TTGCCCCAGC GCTAGCGAAA CCCACGTCAC
TACTTCTCTC CCGAGACGAC ACAGGACGAC GACACACCTC GTCAGAAGCA AAGCGGGTCC CGATCGCTTT GGTGCGAGTG

+3 G G S A G H T V S G F V S L L A P G A K Q N V Q L I
2081 CGGGGAAGT GCGGCCACA CTGTGTCTGG ATTGTTAGC CTCCTCGCAC CAGGCGCCAA GCAGAACGTC CAGCTGATCA
GCCCCCTTCA CGGCCGGTGT GACACAGACC TAAACAATCG GAGGAGCGTG GTCCGCGGTT CGTCTTGCGG GTCGACTAGT

+3 N T N G S W H L N S T A L N C N D S L N T G W L A G L
2161 ACACCAACGG CAGTTGGCAC CTCATAGCA CGGCCCTGAA CTGCAATGAT AGCCTCAACA CCGCTGGTTT GGCAGGCTT
TGTTGGTTGCC GTCAACCGTG GAGTTATCGT GCCGGGACTT GACGTTACTA TCGGAGTTGT GGCCGACCAA CCGTCCCGAA

+3 F Y H H K F N S S G C P E R L A S C R P L T D F D Q G
2241 TTCTATCACC ACAAGTTCAA CTCTTCAGGC TGTCTGAGA GGCTAGCCAG CTGCCGACCC CTTACCGATT TTGACCAGGG
AAGATAGTGG TGTCAAGTT GAGAAGTCCG ACAGGACTCT CCGATCGGTC GACGGCTGGG GAATGGCTAA AACTGGTCCC

```

FIG. 4D

```

+3 W G P I S Y A N G S G P D Q R P Y C W H Y P P K P C
2321 CTGGGGCCT ATCAGTTATG CCAACGGAAG CGGCCCGGAC CAGGGCCCTT ACTGCTGGCA CTACCCCCCA AAACCTTGCG
GACCCCGGGA TAGTCAATAC GGTGCTTC GCCGGGCTG GTCGGGGGA TGACGACCGT GATGGGGGT TTTGGAACGC

+3 G I V P A K S V C G P V Y C F T P S P V V V G T T D R
2401 GTATTGTGCC CGCGAAGAGT GTGTGTGGTC CGGTATATTG CTTCACTCCC AGCCCCGTGG TGGTGGGAAC GACCGACAGG
CATAACACGG GCGCTTCTCA CACACACCAG GCCATATAAC GAATGAGGG TCGGGGCACC ACCACCTTG CTGGCTGTCC

+3 S G A P T Y S W G E N D T D V F V L N N T R P P L G N
2481 TCGGGCGGC CCACCTACAG CTGGGTGAA AATGATACGG ACGTCTTCGT CCTTAACAAT ACCAGGCCAC CGCTGGGCAA
AGCCCGCGCG GGTGGATGTC GACCCCACTT TTAATATGCC TGCAGAAGCA GGAATTGTTA TGGTCCGGTG GCGACCCGTT

+3 W F G C T W M N S T G F T K V C G A P P C V I G G A
2561 TTGGTTCGGT TGTACCTGGA TGAATCAAC TGGATTACCC AAAGTGTGCG GAGCGCCTCC TTGTGTCATC GGAGGGGCGG
AACCAGCCA ACATGGACCT ACTTGAGTGG ACCTAAGTGG TTTACACGC CTCGGGAGG AACACAGTAG CCTCCCCGCC

+3 G N N T L H C P T D C F R K H P D A T Y S R C G S G P
2641 GCAACACAC CCTGCACTGC CCCACTGATT GCTTCCGCAA GCATCCGAC GCACATACT CTCGGTGGG CTCCGGTCCC
CGTTGTGTG GGACGTGACG GGTGACTAA CGAAGCGGT CGTAGGCTG CGGTGTATGA GAGCCACGCC GAGGCCAGGG

+3 W I T P R C L V D Y P Y R L W H Y P C T I N Y T I F K
2721 TGGATCACAC CCAGTGCCT GGTGACTAC CCGTATAGG TTTGGCATT TCTGTATACC ATCAACTACA CCATATTTAA
ACCTAGTGTG GGTCCACGGA CCAGCTGATG GGCATATCCG AAACCGTAAT AGGAACATGG TAGTTGATGT GGTATAAATT

+3 I R M Y V G G V E H R L E A A C N W T R G E R C D L
2801 AATCAGGATG TACGTGGGAG GGTTCGAACA CAGGTGGAA GCTGCCTGCA ACTGGACGCG GGGCGAACGT TGGATCTGG
TTAGTCTTAC ATGCACCTC CCCAGTTGT GTCCGACCTT CGACGGACGT TGACCTGGC CCCGCTTGA ACGTAGACC

+3 E D R D R S E I D M E N I T S G F L G P L L V L Q A G
          ClaI
          -----
2881 AAGATAGGA CAGTCCGAG ATCGATATGG AGAATATCAC ATCAGGATTC CTAGGACCCC TGCTCGTGT ACAGGGGGG
TTCTATCCCT GTCCAGGCTC TAGCTATACC TCTTGTAGTG TAGTCTTAAG GATCCTGGG ACAGACACAA TGTCCGCCCC

```

FIG. 4E

```

+3 F F L L T R I L T I P Q S L D S W W T S L N F L G G S
2961 TTTTCTTGT TGACAAAGAT CCTCACAATA CCGCAGAGTC TAGACTCGTG GTGGACTTCT CTCAATTTTC TAGGGGATC
    AAAAGAACA ACTGTTCTTA GGAGTGTTAT GCGCTCTCAG ATCTGAGCAC CACCTGAAGA GAGTTAAAG ATCCCCCTAG

+3 P V C L G Q N S Q S P T S N H S P T S C P P I C P G
3041 TCCCGTGTGT CTGGGCCAAA ATTGCAGTC CCCAACCTCC AATCACTCAC CAACCTCCTG TCCTCCAATT TGTCCTGGTT
    AGGCACACA GAACCGGTTT TAAGCGTCAG GGGTTGGAGG TTAGTGAGTG GTTGGAGGAC AGGAGGTAA ACAGGACCAA

+3 Y R W M C L R R F I I F L F I L L L C L I F L L V L L
3121 ATCGCTGGAT GTGTCTGGG CGTTTTATCA TATTCCTCTT CATCCTGCTG CTATGCCCTCA TCTTCTTATT GGTTCTTCTG
    TAGGCACCTA CACAGACGCC GCAAATAGT ATAAGGAGAA GTAGGACGAC GATACGGAGT AGAAGAATAA CCAAGAGAGAC

+3 D Y Q G M L P V C P L I P G S T T T S T G P C K T C T
3201 GATTATCAAG GTATGTTGCC CGTTGTCTCT TAAATTCAG GATCAACAAC AACCAGTACG GGACCATGCA AAACCTGCAC
    CTAATAGTTC CATAACAACG GCAAACAGGA GATTAAGGTC CTAGTTGTTG TTGGTCATGC CCTGGTACGT TTTGGACGTC

+3 T P A Q G N S M F P S C C C T K P T D G N C T C I P
3281 GACTCCTGCT CAAGGCAACT CTATGTTTCC CTCATGTTGC TGTACAAAC CTACGGATGG AAATTGCACC TGTATTCCCA
    CTGAGGACGA GTTCCGTTGA GATACAAAG GAGTACAACG ACATGTTTGG GATGCCCTACC TTTAACGTGG ACATAAGGGT

+3 I P S S W A F A K Y L W E W A S V R F S W L S L L V P
3361 TCCCATCGTC CTGGGCTTTC GCAAATACC TATGGAGTG GGCCTCAGTC CGTTTCTCTT GGCTCAGTTT ACTAGTGCCA
    AGGGTAGCAG GACCCGAAAG CGTTTATGG ATACCTCAG CCGGAGTCAG GCAAAGAGAA CCGAGTCAAA TGATCACGGT

+3 F V Q W F V G L S P T V W L S A I W M M W Y W G P S L
3441 TTGTTCAGT GGTTCGTAGG GCTTTCCTCC ACTGTTTGGC TTTCAGCTAT ATGGATGATG TGGTATTGGG GGCCAAGTCT
    AAACAAGTCA CCAAGCATCC CGAAAGGGGG TGACAAACCG AAAGTCGATA TACCTACTAC ACCATAACCC CCGGTTTCAGA

+3 Y S I V S P F I P L L P I F F C L W V Y I *
    EcorI
    -----
3521 GTACAGCATC GTGAGTCCCT TTATACCGCT GTTACCAATT TTCTTTTGTG TCTGGGTATA CATTTAAGAA TTCAGACTCG
    CATGTCGTAG CACTCAGGGA AATATGGCGA CAATGGTTAA AAGAAACACAG AGACCCCATAT GTAAATTTCTT AAGTCTGAGC

```

FIG. 4F

BamHI

```

-----
3601 AGCAAGTCTA GAAAGGCGG CCAAGATATC AAGGATCCAC TACGCGTTAG AGCTCGCTGA TCAGCCTCGA CTGTGCCTTC
    TCGTTCAGAT CTTTCCGCGC GGTTCCTATAG TTCCTAGTG ATCGCAATC TCAGCGACT AGTCGAGCT GACACGGAAG

3681 TAGTTGCCAG CCATCTGTTG TTTGCCCTC CCCCGTGCCT TCCTTGACCC TGAAGGTGC CACTCCCACT GTCCTTTCCT
    ATCAACGGTC GGTAGACAAC AAACGGGAG GGGGCACGGA AGGAACCTGG ACCTTCCAG GTGAGGGTGA CAGGAAGGA

3761 AATAAAATGA GGAATTTGCA TCGCATTTGC TGAGTAGGTG TCATTCTATT CTGGGGGTG GGTGGGGCA GGACAGCAAG
    TTATTTTACT CCTTTAACGT AGCGTAACAG ACTCATCCAC AGTAAGATAA GACCCCCAC CCCACCCCGT CCTGTCTGTC

3841 GGGGAGGATT GGAAGACAA TAGCAGGCAT GCTGGGAGC TCTTCCGCTT CCTCGCTCAC TGACTCGCTG CGCTCGGTG
    CCCCTCCTAA CCTTCTGTT ATCGTCCGTA CGACCCCTCG AGAAGCGAA GGAGCGAGTG ACTGAGCGAC GCGAGCCAGC

3921 TTCGGCTGG GCGAGCGGTA TCAGCTCACT CAAAGCGGT AATACGGTTA TCCACAGAAT CAGGGGATAA CGCAGGAAAG
    AAGCCGACG CGCTCGCCAT AGTCGAGTGA GTTTCGCCA TTATGCCAAT AGGTGCTTA GTCCCCATT GGTCTCTTTC

4001 AACATGTGAG CAAAAGGCA GCAAAAGCC AGGAACCGTA AAAAGGCCG GTTCTGGCG TTTTTCATA GGTCCGCCC
    TTGTACACT GTTTTCCGGT CGTTTCCGG TCCTTGGCAT TTTTCCGGC CAACGACCG AAAAAGGTAT CCGAGGCGGG

4081 CCCTGACGAG CATCACAAA ATCGACGCTC AAGTCAGAG TGGCGAAACC CGACAGGACT ATAAAGATAC CAGGCGTTTC
    GGGACTGCTC GTAGTGTTT TAGCTGGAG TTCAGTCTC ACCGCTTGG GCTGTCTGA TATTCTATG GTCCGCAAAG

4161 CCCCTGGAAG CTCCTCTGCG CGCTCTCCTG TTCCGACCTT GCCGTTACC GGATACCTGT CCGCCTTTCT CCCTTCGGGA
    GGGGACCTTC GAGGGAGCAC GCGAGAGGAC AAGGCTGGGA CCGCGAATGG CCTATGGACA GCGCGAAAGA GGAAGCCCT

4241 AGCGTGGCG TTTCTCAATG CTCACGCTGT AGGTATCTCA GTTCGGTGTG GGTCTGTCG TCCAAGCTGG GCTGTGTGCA
    TCGCACCCGG AAAGAGTTAC GAGTGGACA TCCATAGAT CAAGCCACAT CCAGCAAGCG AGGTTCGACC CGACACACGT

4321 CGAACCCCCC GTTCAGCCCG ACCGCTGCGC CTTATCCGGT AACTATCGTC TTGAGTCCAA CCCGGTAAGA CAGACTTAT
    GCTTGGGGG CAAGTCGGG TGGCGACGG GAATAGGCCA TTGATAGCAG AACTCAGGT GGGCCATTCT GTGCTGAATA

4401 CGCCACTGGC AGCAGCCACT GGTAAACAGGA TTAGCAGAGC GAGGTATGTA GCGCGTGTCTA CAGAGTTCTT GAAGTGGTGG
    GCGGTGACCG TCGTGGGTGA CCATTGTCTT AATCGTCTCG CTCATACAT CCGCCACGAT GTCTCAAGAA CTTACCCACC
  
```

FIG. 4G

4481 CCTAACTACG GCTACACTAG AAGGACAGTA TTTTGGTATCT GCGCTCTGCT GAAGCCAGTT ACCTTCGGAA AAAGAGTTGG  
GGATTGATGC CGATGTGATC TTCTCTGTCAT AAACCATAGA CGCGAGACGA CTTCGGTCAA TGGGAAGCCCTT TTCTCAACC

4561 TAGCTCTTGA TCCGGCAAAAC AAACCAACCGC TGGTAGCGGT GGTCTTTTGG TTTGCAAGCA GCAGATTACG CGCAGAAAAA  
ATCGAGAACT AGGCCGTTTG TTTGGTGGCG ACCATCGCCA CCAAAAAAAC AAACGTTTCGT CGTCTAATGC GCGTCTTTTT

4641 AAGGATCTCA AGAAGATCCT TTGATCTTTT CTACGGGGTC TGACGCTCAG TGGAAACGAAA ACTCACGTTA AGGGATTTTG  
TTCTTAGAGT TCCTTAGGA AACTAGAAAA GATGCCCCAG ACTGGAGTC ACCTTGCTTT TGAGTGCAAT TCCCTAAAAAC

4721 GTCATGAGAT TATCAAAAAG GATCTTCACC TAGATCTTTT TAAATTAATAA ATGAAGTTTT AAATCAATCT AAAGTATATA  
CAGTACTCTA ATAGTTTTTC CTAGAAGTGG ATCTAGGAAA ATTTAATTTT TACTTCAAAA TTAGTTAGA TTTCATATAT

4801 TGAGTAAACT TGGTCTGACA GTTACCAATG CTTAATCAGT GAGGCACCTA TCTCAGCGAT CTGTCTATTT CGTTCATCCA  
ACTCATTGA ACCAGACTGT CAATGGTTAC GAATTAGTCA CTCGGTGGAT AGAGTCGCTA GACAGATAAA GCAAGTAGGT

4881 TAGTTGCCG ACTCCCGTC GTGTAGATAA CTACGATACG GGAGGGCTTA CCATCTGGCC CCAGTGCTGC AATGATACCG  
ATCAACGAC TGAGGGGAC CACATCTATT GATGTATGC CCTCCCGAAT GGTACACGCG GGTACACGACG TTAATATGGC

4961 CGAGACCCAC GCTACCGGC TCCAGATTTA TCAGCAATAA ACCAGCCAGC CGGAAGGGCC GAGCGCAGAA GTGGTCCTGC  
GCTCTGGGTG CGAGTGGCCG AGGTCTAAT AGTCTTATT TGGTGGTGC GCCTTCCCG CTGCGGTCCT CACCAGGACG

5041 AACTTTATCC GCCTCCATCC AGTCTATTAA TTGTTGCCGG GAAGCTAGAG TAAGTAGTTC GCCAGTTAAT AGTTTGCGCA  
TTGAAATAGG CGGAGGTAGG TCAGATAATT AACAAACGGCC CTTCGATCTC ATTCATCAAG CGGTCAATTA TCAAAACGCGT

5121 ACGTTGTGC CATTGCTACA GGCATCGTGG TGTACAGCTC GTCGTTTGGT ATGGCTTCAT TCAGCTCCGG TTCCCAACGA  
TGCAACACAG GTAACGATGT CCGTAGCACC ACAGTGGAG CAGCAACCA TACCGAAGTA AGTCGAGGCC AAGGTTGCT

5201 TCAAGGCGAG TTACATGATC CCCCATGTTG TGCAAAAAAG CGGTTAGCTC CTTCGGTCCCT CCGATCGTTG TCAGAAGTAA  
AGTTCCGCTC AATGTACTAG GGGGTACAAC ACGTTTTTC GCCAATCGAG GAAAGCCAGGA GGCTAGCAAC AGTCTTCATT

5281 GTTGGCCGCA GTGTTATCAC TCATGTTAT GGCAGCACTG CATTAATTCTC TTACTGTCTAT GCCATCCGTA AGATGCTTTT  
CAACCGGGGT CACATAGTG AGTACCAATA CCGTCGTGAC GTATTAGAG AATGACAGTA CCGTAGGCAT TCTACGAAAA

FIG. 4H



5361 CTGTGACTGG TGAGTACTCA ACCAAGTCAT TCTGAGAATA GTGTATGCGG CGACCGAGTT GCTCTTGCCC GCGTCAATA  
GACACTGACC ACTCATGAGT TGGTTCAGTA AGACTCTTAT CACATACGCC GCTGGCTCAA CGAGAACGGG CCGCAGTTAT

5441 CCGGATAATA CCGCGCCACA TAGCAGAACT TTAAAAGTGC TCATCAATTGG AAAACGTTCT TCGGGGCGAA AACTCTCAAG  
GCCCTATTAT GCGCGGTGT ATCGTCTTGA AATTTTCACG AGTAGTAACC TTTTGCAAGA AGCCCCGCTT TTGAGAGTTC

5521 GATCTTACCG CTGTTGAGAT CCAGTTCGAT GTAACCCACT CGTGCACCCA ACTGATCTTC AGCATCTTTT ACTTTCACCA  
CTAGAATGGC GACAACTCTA GGTCAGCTA CATTGGGTGA GCACGTGGGT TGA CTAGAAG TCGTAGAAAA TGAAAGTGGT

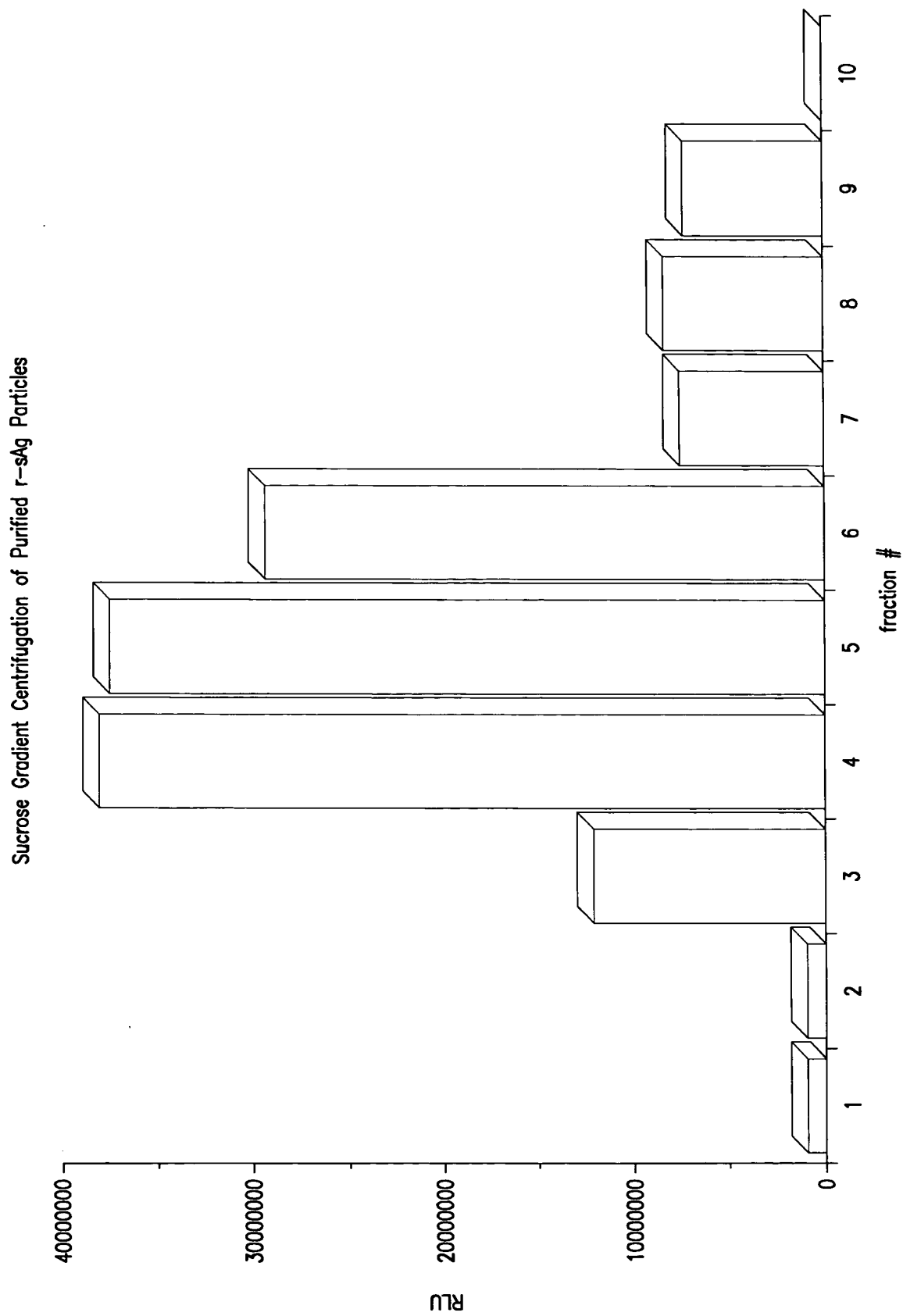
5601 GCGTTTCTGG GTGAGCAAAA ACAGGAAGGC AAAATGCCGC AAAAAAGGGA ATAAGGGCGA CACGGAAAATG TTGAATACTC  
CGCAAAGACC CACTCGTTT TGTCTTCCG TTTTACGGCG TTTTTCCT TATTCGCCCT GTGCCCTTAC AACTTATGAG

5681 ATACTCTTCC TTTTTCATA TTATTGAAGC ATTATCAGG GTTATTGTCT CATGAGCGGA TACATATTTG AATGTATTTA  
TATGAGAAGG AAAAAAGTTAT AATAACTTCG TAAATAGTCC CAATAACAGA GTACTCGCCT ATGTATAAAC TTACATAAAT

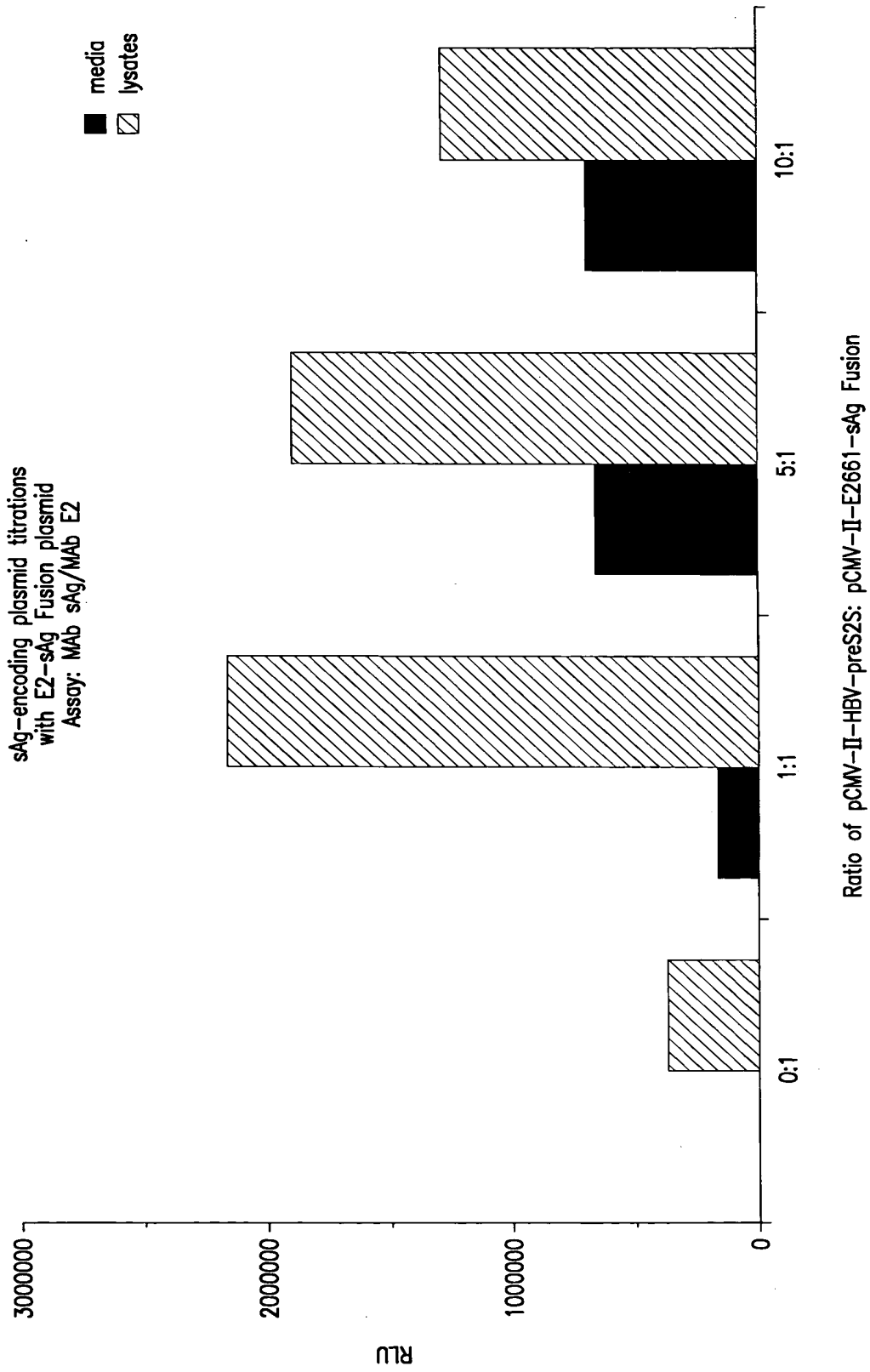
5761 GAAAAATAA CAAATAGGGG TTCCGCGCAC ATTTCCCGA AAAGTGCCAC CTGACGTCTA AGAAACCATT ATTATCATGA  
CTTTTATTI GTTTATCCCC AAGGCGCGTG TAAAGGGGCT TTTACCGGTG GACTGCAGAT TCTTTGGTAA TAATAGTACT

5841 CATTAACTTA TAAAAATAGG CGTATCACGA GGCCCTTTTCG TC  
GTAATTGGAT ATTTTATCC GCATAGTGCT CCGGAAAGC AG

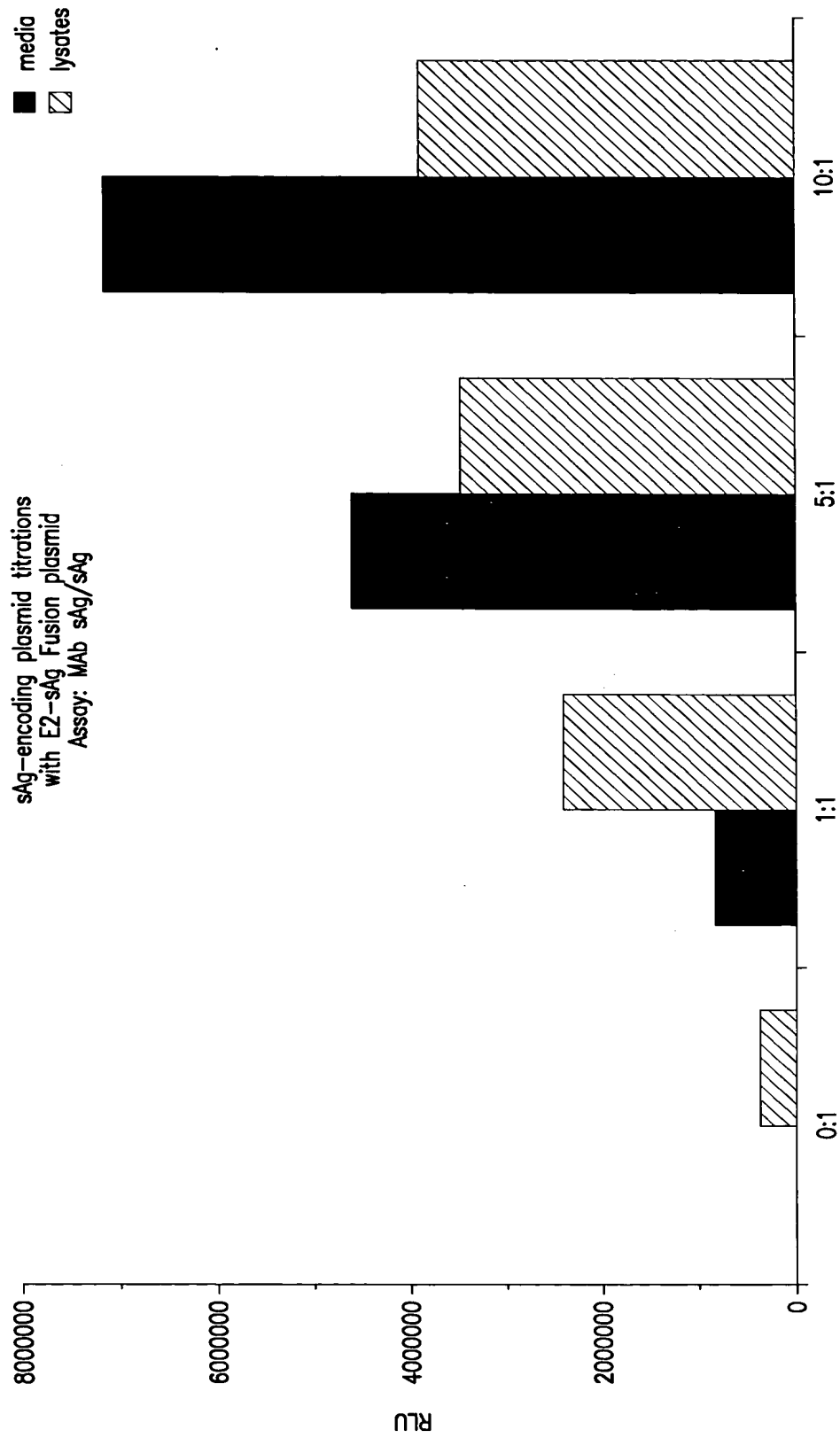
FIG. 4I



**FIG. 5**



**FIG. 6A**



**FIG. 6B**

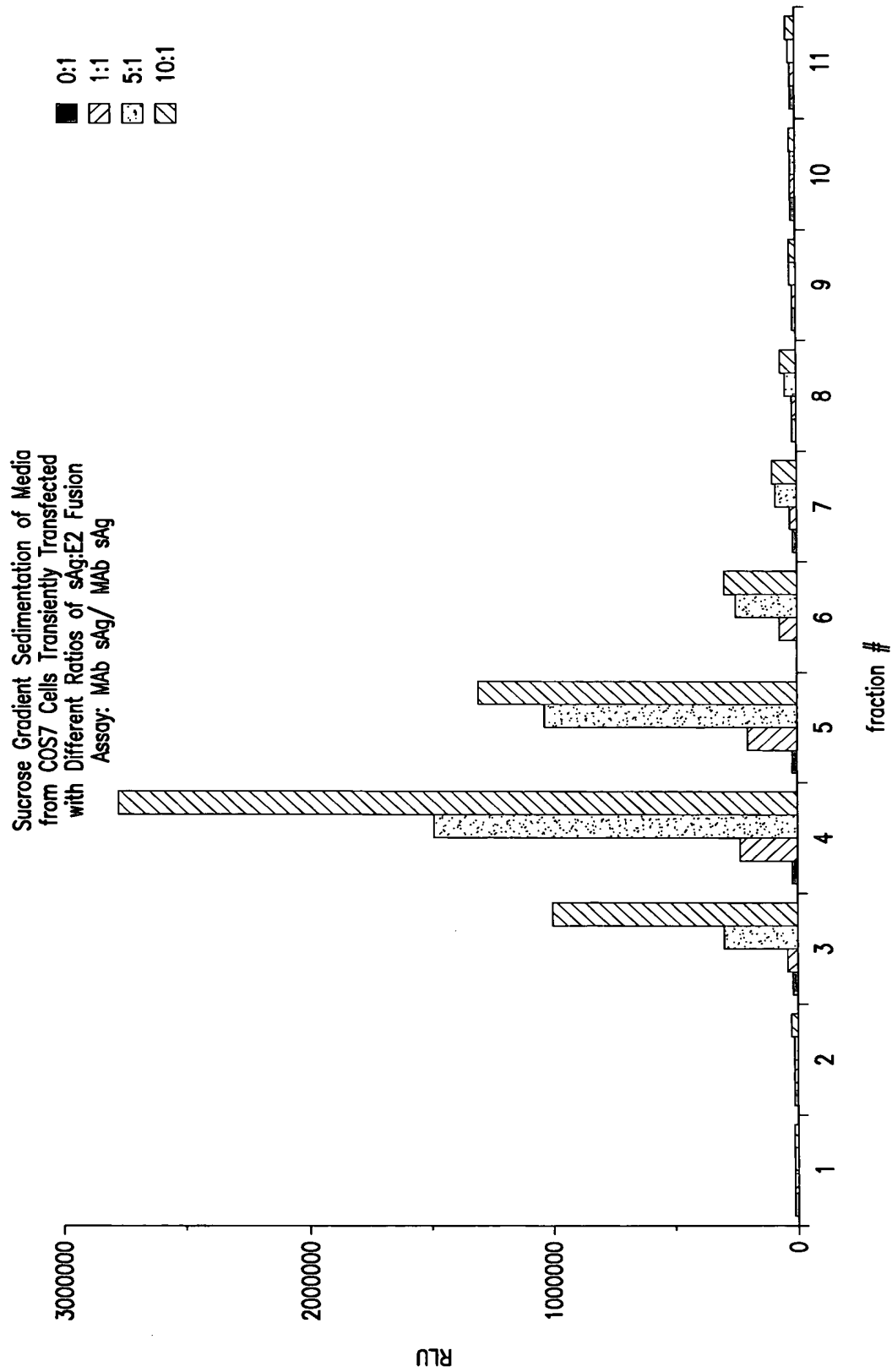


FIG. 7A

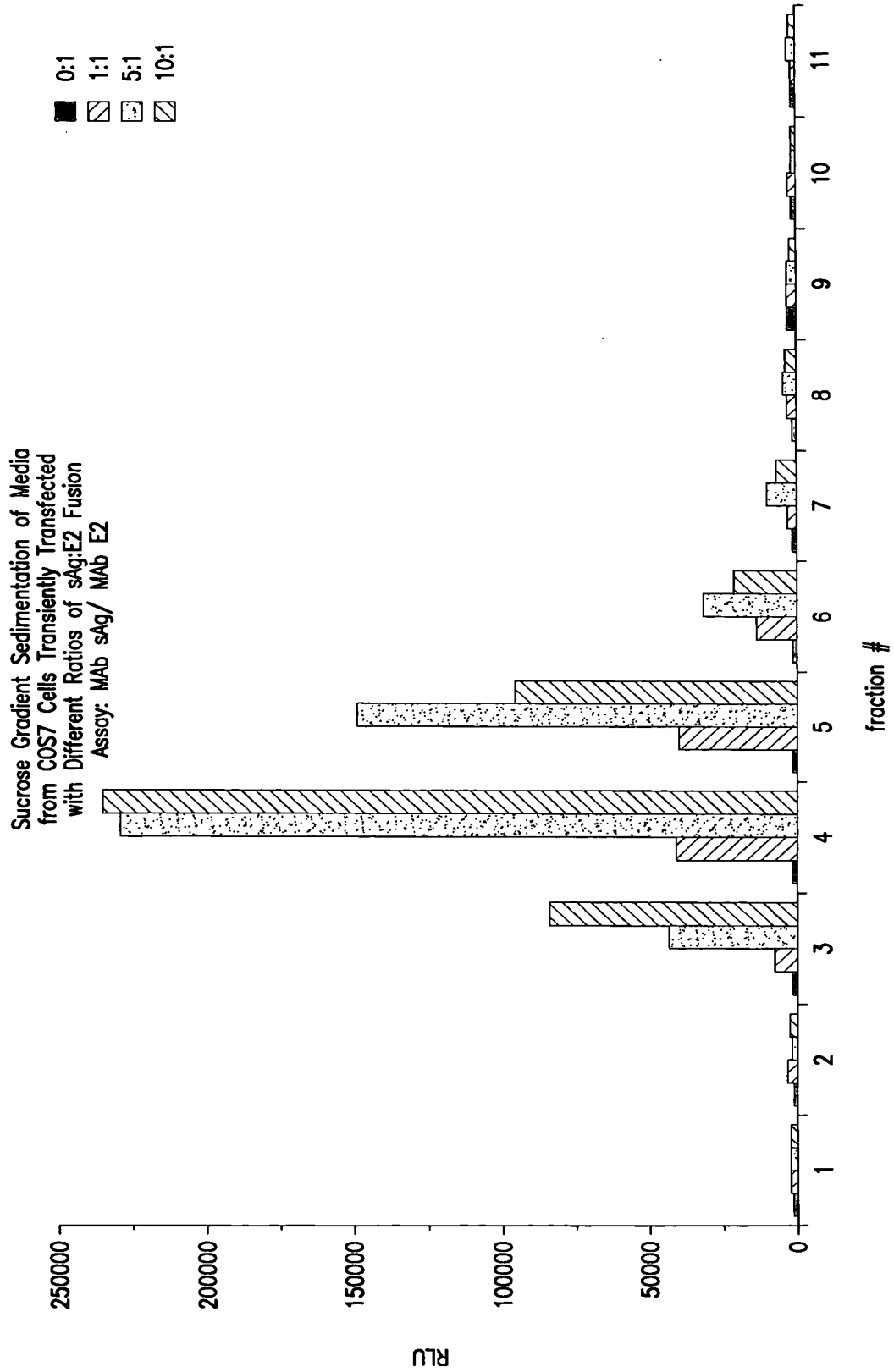


FIG. 7B

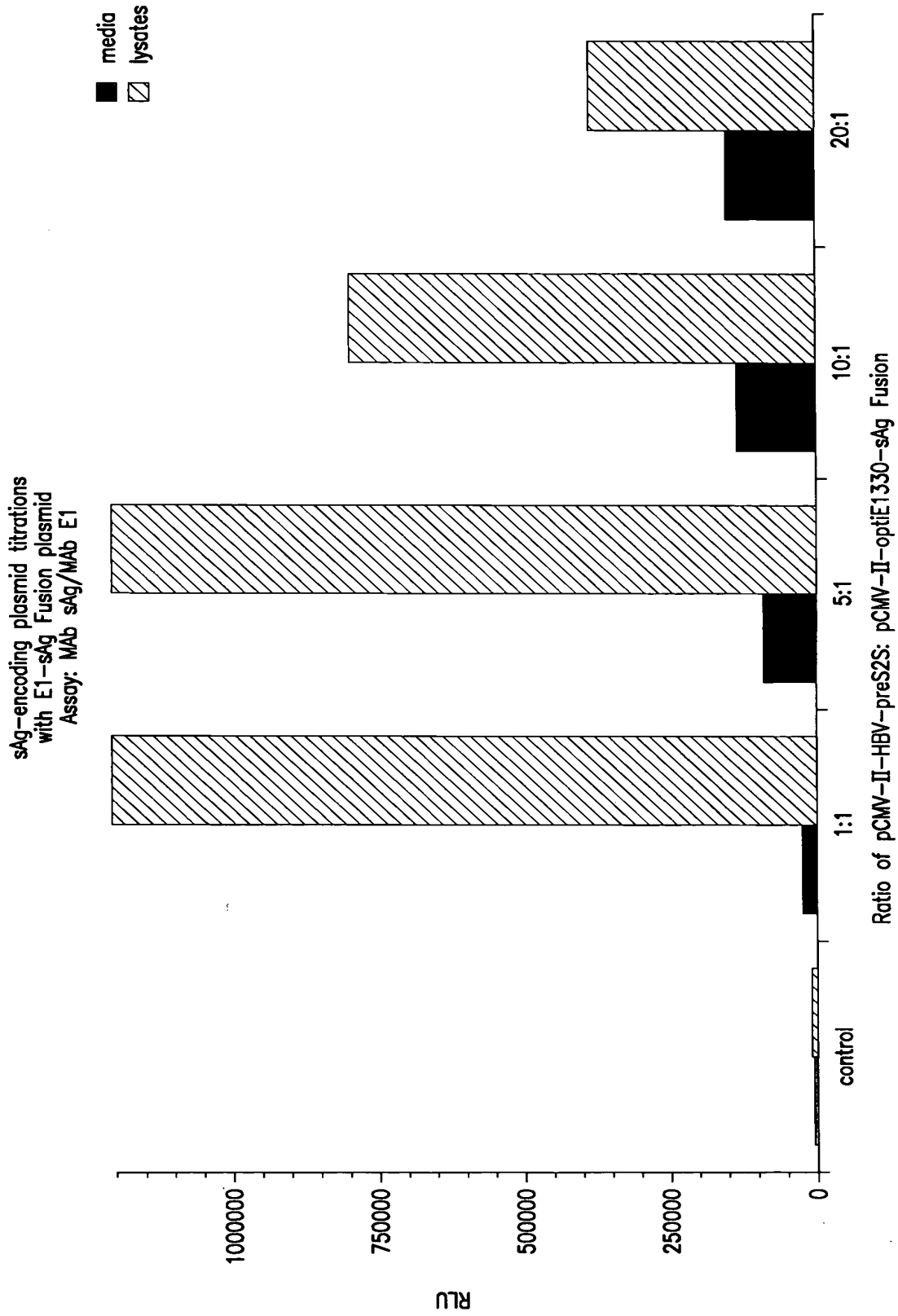


FIG. 8A

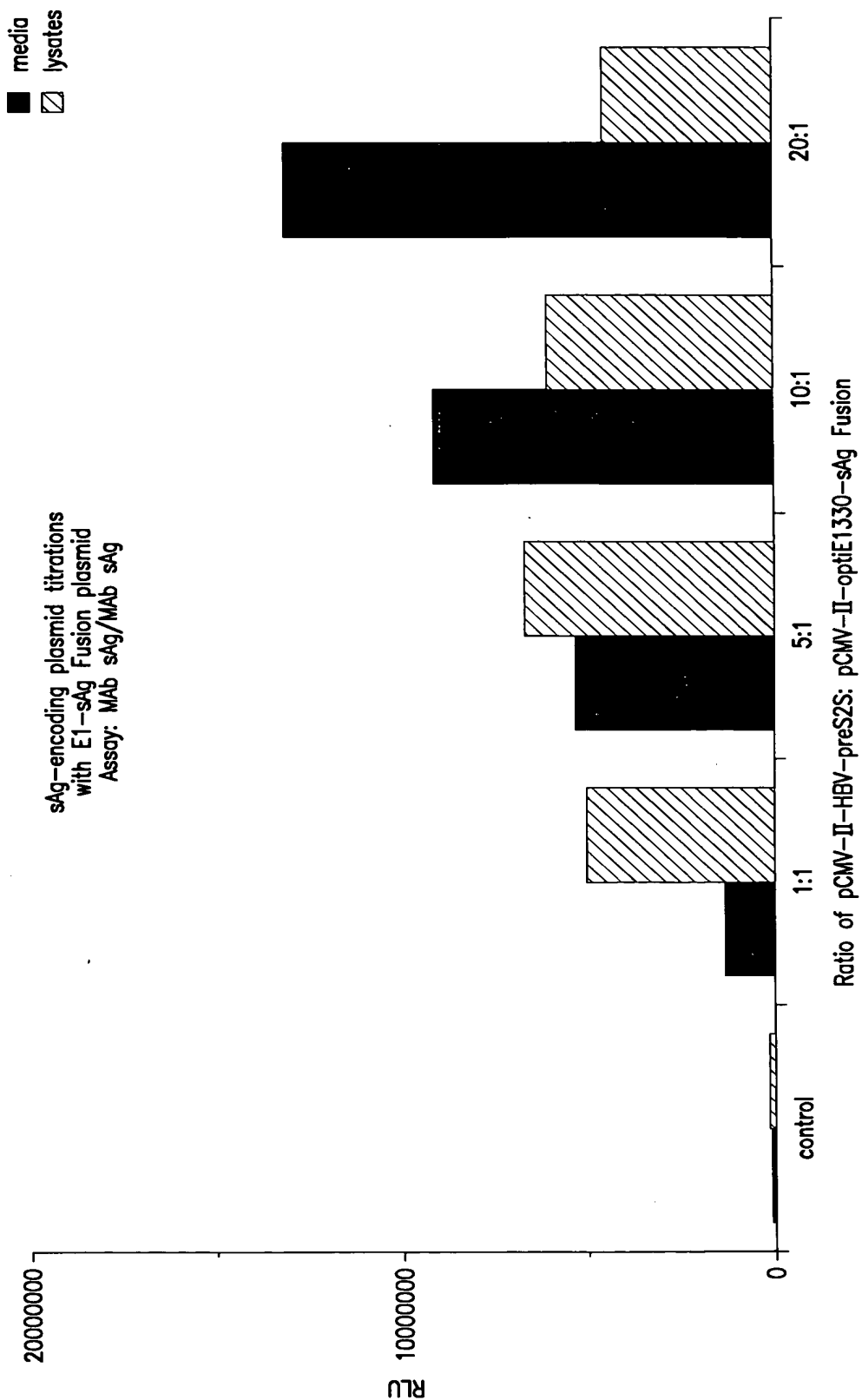
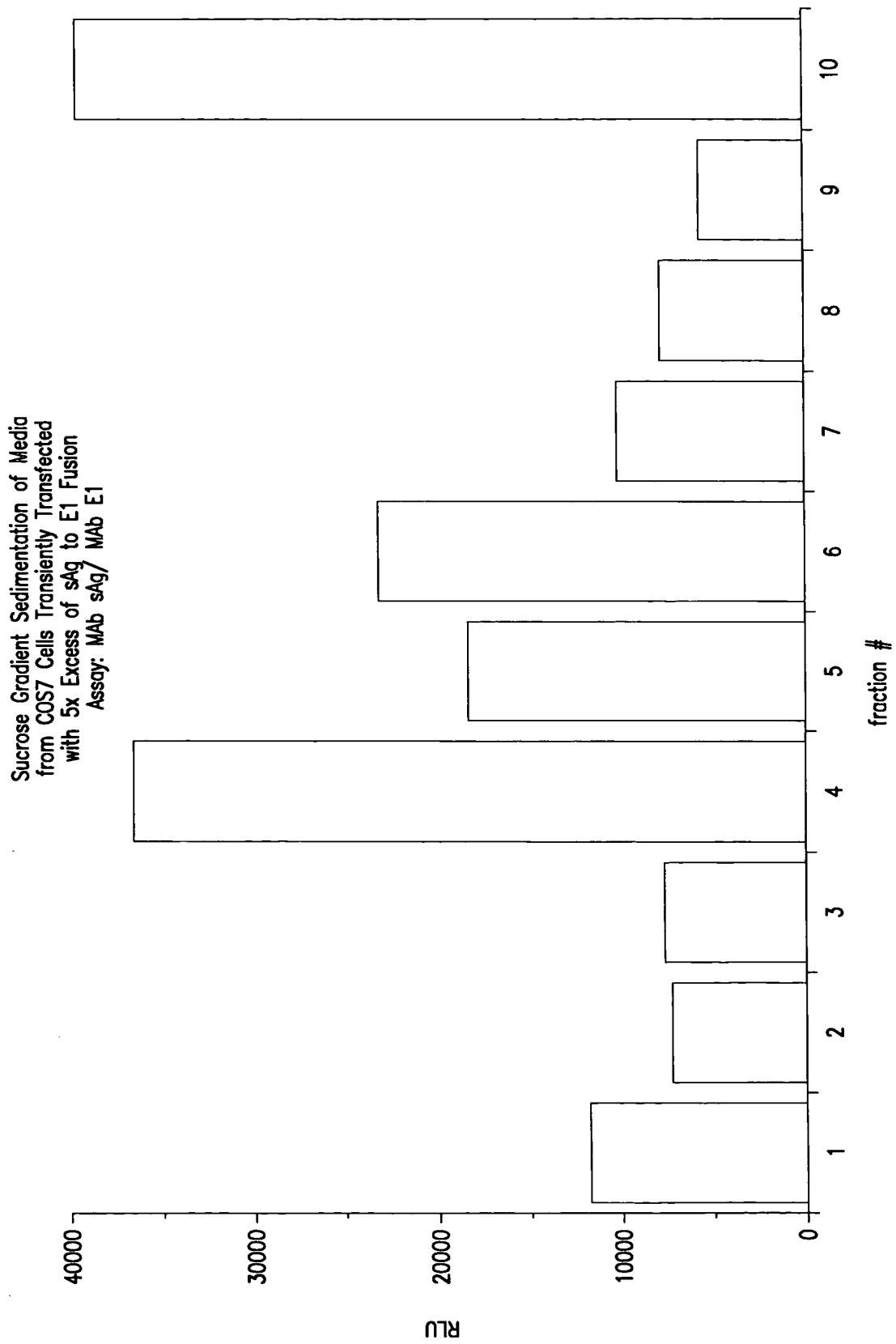
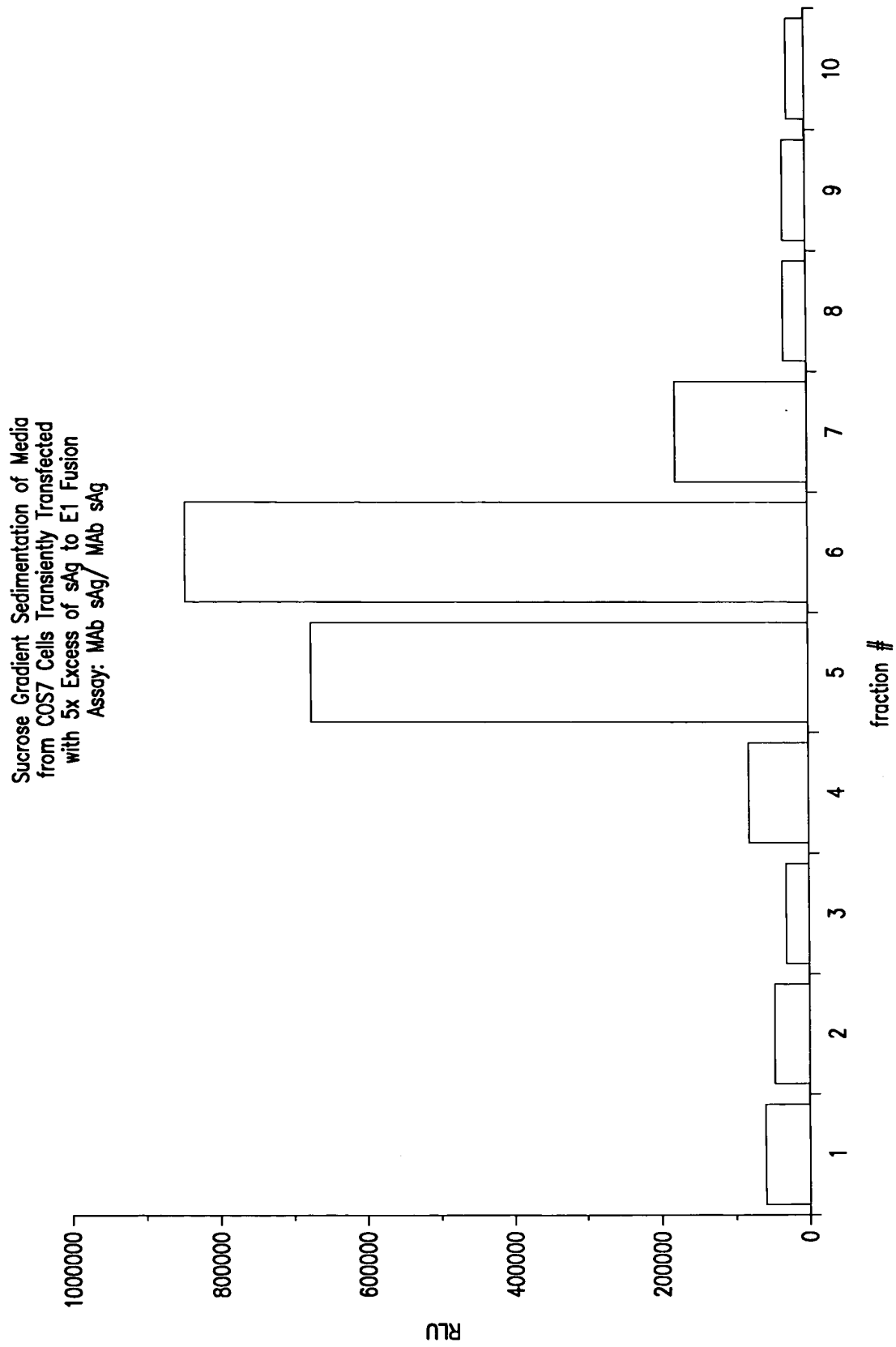


FIG. 8B

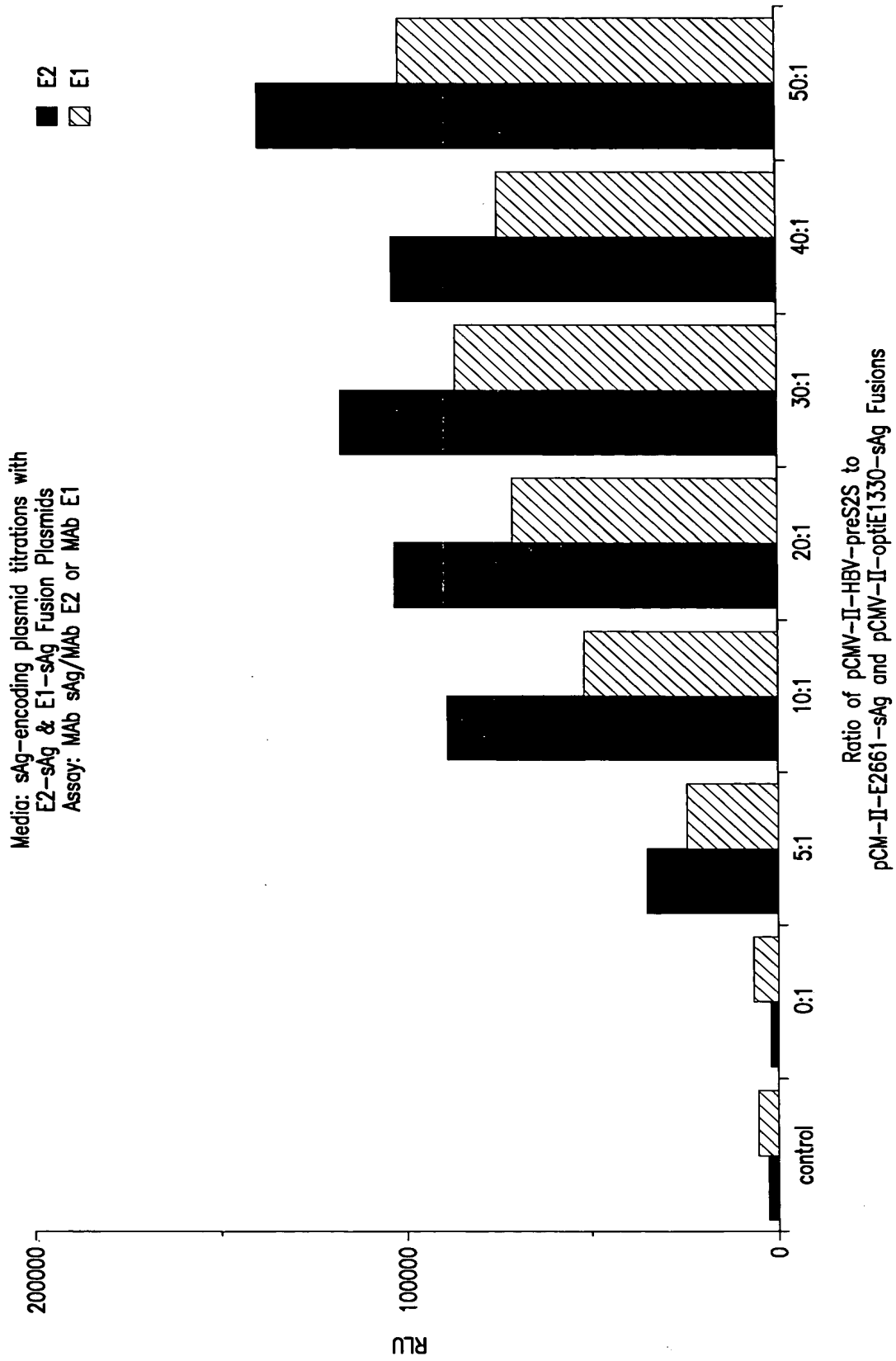




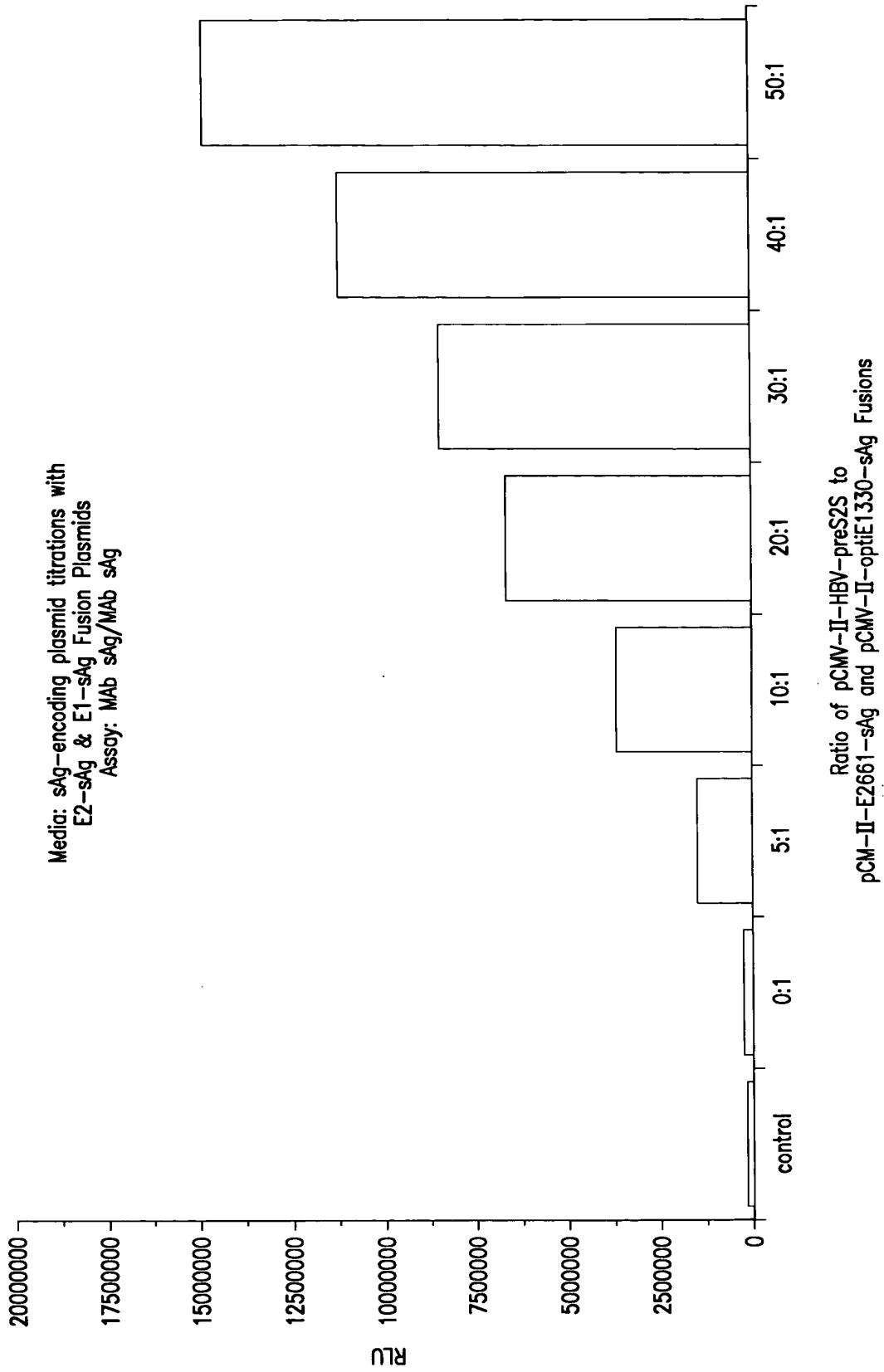
**FIG. 9A**



**FIG. 9B**



**FIG. 10A**



**FIG. 10B**